MADField: Multi-fidelity Amortized Density Field for Adsorption in Nanoporous Materials

Anonymous Authors[1]

Abstract

High-throughput computational screening of nanoporous materials for gas storage and separation requires fast and accurate characterization of adsorption equilibrium. Particle-based grand canonical Monte Carlo (GCMC) and density-based classical density functional theory (cDFT) provide simulation-based estimates of gas uptake and adsorbate density fields, but their accuracy–speed tradeoff remains insufficient for large-scale screening. In this work, we address this gap with Multi-fidelity Amortized Density Field (MADField) that reframes adsorption prediction as equilibrium density-field estimation. MADField exploits density supervision from complementary fidelity levels: cDFT provides broad, scalable density supervision, while particle-sampling GCMC provides higher-fidelity density labels after coarse graining. To enable this, we generate and release a large-scale cDFT adsorption dataset spanning thousands of MOF frameworks and nine adsorbates. We first train a density predictor on cDFT labels, use its predictions to initialize previously failed cDFT calculations and expand the cDFT label set, and then fine-tune the model on GCMC-derived density fields. Across a comprehensive benchmark spanning multiple adsorbates and material classes, MADField improves uptake accuracy over the strongest baselines by 6.0× and 15.4× for cDFT and GCMC approximation, respectively. Its predicted density fields also accelerate cDFT solvers, reducing number of solver steps by 2.0× and recovering 42% failures of standard cDFT. We will release the underlying cDFT benchmark dataset of 280,000 calculations, generated with 3,600 H200 GPU hours, to support future work.

[1]Anonymous Institution, Anonymous City, Anonymous Region, Anonymous Country. Correspondence to: Anonymous Author <anon.email@domain.com>.

Submitted to the AI for Science workshop (ICML 2026). Do not distribute.

1. Introduction

High-throughput computational screening is central to discovering nanoporous materials for gas storage and separation (Wilmer et al., 2012). Metal–organic frameworks (MOFs) and amorphous carbons (ACs), for example, have been widely screened for methane storage (Mason et al., 2014) and carbon capture (Lin et al., 2012; Charalambous et al., 2024). The primary screening observable is gas uptake, i.e., the amount adsorbed at a given pressure and temperature. Computing uptake, however, requires characterizing adsorption equilibrium: the distribution of gas configurations inside a complex porous material under a specified thermodynamic condition.

Adsorption equilibrium can be modeled through two complementary views. *Grand canonical Monte Carlo (GCMC)* provides a *particle-based view*, sampling gas molecules in the grand canonical ensemble through insertion, deletion, and roto-translational moves as illustrated in Figure 1(a) (Adams, 1975). As it estimates ensemble averages from particle configurations under a chosen force field, GCMC is widely used as a reference for adsorption screening (Dubbeldam et al., 2016; Ran et al., 2024). However, long Markov chains with millions of Monte Carlo steps limit its use in high-throughput settings. *Classical density functional theory (cDFT)* provides a complementary *density-based view*, replacing stochastic sampling with optimization of a grand-potential functional over the one-body density field as illustrated in Figure 1(b) (Evans, 1979; Roth, 2010). Recent work shows that cDFT can reproduce GCMC-level adsorption trends in MOFs at lower cost (Dufour-Décieux et al., 2025; Thiele et al., 2026). Nevertheless, cDFT relies on an approximate free-energy functional, and each material–adsorbate–condition still requires an iterative solve.

Existing machine learning approaches do not fully exploit this particle–density structure. Most methods regress scalar uptake from structural, textural, or energetic descriptors (Sarikas et al., 2024; Chen et al., 2022), but discard where gas adsorbs inside the material and leave a non-trivial accuracy gap relative to simulation references. A smaller line of work predicts adsorbate density fields (Sun & Siepmann, 2024; Burner et al., 2025), but existing models either produce normalized probability-like fields that integrate to

unity, or require the uptake itself as an auxiliary input. Thus, they cannot independently recover the equilibrium adsorbate density.

Contribution. In this work, we introduce Multi-fidelity Amortized Density Field (MADField) for adsorption simulation in nanoporous materials. Given a porous framework, adsorbate, and thermodynamic condition, MADField predicts the unnormalized equilibrium adsorbate density; scalar uptake is then recovered by integration rather than by an unconstrained regression head. This density-level formulation provides a unified target for screening: it predicts gas uptake, preserves spatial adsorption structure, and can initialize iterative cDFT solvers.

The key challenge is that density supervision is available at different fidelity-cost regimes. cDFT provides relatively inexpensive density fields but inherits the approximation error of its free-energy functional, while GCMC provides costly particle-sampling references that can be coarse-grained into density. We therefore pre-train MADField on large-scale cDFT density data and adapt it to sparse GCMC-derived density fields, learning a broad density prior from cDFT and correcting the remaining fidelity gap toward particle-simulation references as illustrated in Figure 1(c).

Our main contributions:

- **Equilibrium density as a unified target.** We formulate adsorption prediction as unnormalized equilibrium adsorbate density field modeling. MADField predicts the 3D equilibrium density in physical units, so uptake N is obtained by integration while spatial adsorption structure is retained.
- **Multi-fidelity density learning.** We bridge cDFT and GCMC by pre-training MADField on cDFT density fields across nine non-polar adsorbates and fine-tuning it to a GCMC dataset that is $14.7\times$ smaller. This cDFT-to-GCMC transfer outperforms GCMC-only training, especially on out-of-distribution materials, including amorphous carbons (ACs), polymers of intrinsic microporosity (PIMs), hyper-cross-linked polymers (HCPs), and kerogens.
- **Density-based acceleration of cDFT.** Since MADField predicts density fields in physical units, its output can serve as a high-quality initial density for the original cDFT fixed-point solver. Such initialization allows solving 42% of cases where the standard initialization fails and achieves a $2.0\times$ reduction in the number of iterations for the solver to converge.
- **Comprehensive adsorption benchmark.** We benchmark both uptake and density prediction across cDFT and GCMC labels, covering nine adsorbates and diverse material classes. On MOFs, MADField reduces uptake MAE over the strongest baselines by $6.0\times$ for cDFT labels and $15.4\times$ for GCMC labels. On out-of-distribution disordered materials, it also reduces uptake MAE by $4.9\times$ and $3.5\times$ on cDFT and GCMC labels, respectively. We will release the cDFT benchmark dataset of 280,000 calculations with 3,600 H200 GPU hours.

2. Background and related work

This section introduces adsorption equilibrium and its particle- and density-based formulations (Section 2.1), followed by related work on machine learning for adsorption prediction (Section 2.2).

2.1. Adsorption equilibrium as particle and density fields

Adsorption is the process by which guest gas molecules accumulate on the surfaces and within the pores of a host porous material. For a rigid host framework $\mathcal{M}$, guest species s, temperature T, and pressure P, the central field-level object is the equilibrium one-body number density $\rho_{\mathrm{eq}}(\mathbf{r})$, defined as the expected number of adsorbate molecule centers per unit volume at position $\mathbf{r}$ in the periodic unit cell $\mathcal{V}$. *Scalar uptake* N and *adsorption isotherm* $\mathcal{I}_T$ are obtained from this field:

$$\begin{aligned} N(P,T;\mathcal{M},s) &= \int_{\mathcal{V}} \rho_{\mathrm{eq}}(\mathbf{r})\, d\mathbf{r}, \\ \mathcal{I}_T(\mathcal{M},s) &= \{N(P,T;\mathcal{M},s) : P \in \mathcal{P}\}. \end{aligned} \tag{1}$$

Scalar uptake is the primary quantity used to rank porous materials in high-throughput screening, since practical applications often require maximizing the amount of gas stored or captured at a target pressure and temperature. The adsorption isotherm further characterizes how uptake changes with pressure, enabling evaluation of process-relevant quantities such as working capacity between adsorption and desorption pressures, pressure-dependent selectivity, and operating-window robustness.

Particle view: grand canonical Monte Carlo. GCMC simulates adsorption by sampling many-body particle configurations in the grand canonical (μVT) ensemble (Adams, 1975). For adsorbate species s, a configuration with n molecules at positions $\mathbf{r}_{1:n} = (\mathbf{r}_1, \ldots, \mathbf{r}_n)$ in a porous material has energy

$$U_n(\mathbf{r}_{1:n};\mathcal{M},s) = \sum_{i=1}^{n} V_{\mathrm{ext}}(\mathbf{r}_i;\mathcal{M},s) + \sum_{i<j} u_{ss}(\mathbf{r}_i,\mathbf{r}_j), \tag{2}$$

where V_{ext} is the framework–fluid potential and u_{ss} is the fluid–fluid pair interaction. At chemical potential $\mu = \mu(P,T;s)$, a Markov chain proposes insertion, deletion, translation, and rotation moves, accepted by a Metropolis criterion so that the chain samples the μVT distribution over variable-size configurations. Adsorption observables

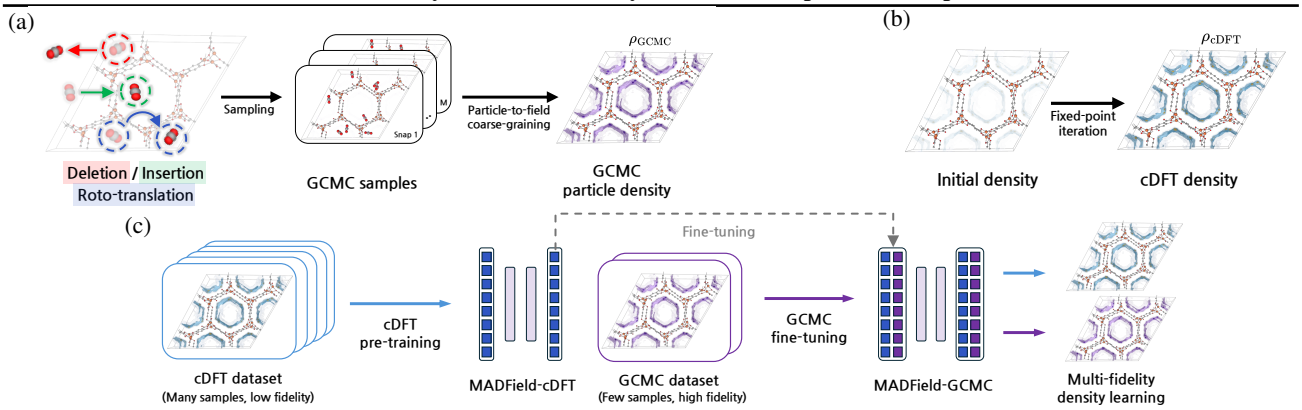

Figure 1. **Multi-fidelity adsorbate density learning from cDFT and GCMC. (a)** GCMC samples a particle ensemble through deletion, insertion, and roto-translation moves, and particle-to-field coarse-graining converts the ensemble into ρ_{GCMC}. **(b)** cDFT solves for ρ_{cDFT} by fixed-point iteration from an initial density. **(c)** MADField is pre-trained on broad cDFT labels and fine-tuned on sparse GCMC labels, combining cDFT coverage with GCMC fidelity.

are then computed as ensemble averages, *i.e.*, uptake is the expected particle count, and the one-body density field is obtained by coarse-graining the particles:

$$N(P,T;\mathcal{M},s) = \langle n \rangle_{\mu,V,T},$$
$$\rho_{\mathrm{eq}}(\mathbf{r}) = \left\langle \sum_{i=1}^{n} \delta(\mathbf{r}-\mathbf{r}_i) \right\rangle_{\mu,V,T}. \tag{3}$$

Integrating this density recovers the uptake consistent with Equation (1). In practice, GCMC with classical force fields is a standard simulation tool for large-scale MOF screening, with studies evaluating over 10^5 of structures for gas storage and separation (Wilmer et al., 2012; Colón & Snurr, 2014; Simon et al., 2015).

Density view: classical density functional theory. cDFT reformulates adsorption equilibrium as a variational problem over the same one-body density field (Evans, 1979; Roth, 2010). For a guest species s at fixed (T,μ), the equilibrium density ρ_{eq} is defined as the minimizer of the grand-potential functional

$$\Omega[\rho;\mathcal{M},s] = F_{\mathrm{id}}[\rho] + F_{\mathrm{exc}}[\rho;s] + \int_{\mathcal{V}} \rho(\mathbf{r})\left[V_{\mathrm{ext}}(\mathbf{r};\mathcal{M},s) - \mu\right] d\mathbf{r}, \tag{4}$$

where F_{id} is the ideal-gas free energy, F_{exc} captures non-ideal fluid correlations, and V_{ext} is the framework–fluid external potential. Setting $\delta\Omega/\delta\rho = 0$ yields the Euler–Lagrange condition

$$\rho_{\mathrm{eq}}(\mathbf{r}) = \rho_{\mathrm{bulk}} \exp\Big[-\beta\, V_{\mathrm{ext}}(\mathbf{r}) - \beta \left.\frac{\delta F_{\mathrm{exc}}}{\delta\rho}\right|_{\rho_{\mathrm{eq}}} + \beta\,\mu_{\mathrm{exc}}\Big], \tag{5}$$

where $\beta = 1/k_B T$, ρ_{bulk} is the bulk fluid density at the given (T,μ), and μ_{exc} is the excess chemical potential of the bulk fluid. Because F_{exc} depends nonlinearly on ρ, this self-consistent equation is solved by fixed-point iteration on a volumetric grid until convergence. Uptake is then computed by integrating the converged density field. Recent work has shown that cDFT can reproduce GCMC-level gas uptake in MOFs at substantially lower computational cost (Dufour-Décieux et al., 2025; Thiele et al., 2026). We use the PC-SAFT functional (Sauer & Gross, 2017), which parameterizes each species by segment number m, segment diameter σ, and dispersion energy ε. Full derivations of the Euler–Lagrange equation, the Boltzmann reference density, the Picard update used in our solver, and the PC-SAFT functional are given in Section A.

2.2. Machine learning for adsorption observables

Machine learning approaches to adsorption have mostly targeted scalar uptake at a given thermodynamic state. These models differ in representation and backbone. Existing models have progressed from building-block and topological embeddings (Lee et al., 2021) to energy-based representations such as 2D energy histograms (Shi et al., 2023) and voxelized potential-energy fields processed by 3D CNNs (Sarikas et al., 2024). More recent work adopts graph and transformer architectures that combine local atomic environments, global textural features, and transferable pre-training across materials and gas species (Chen et al., 2022; Lin et al., 2025; Cui et al., 2023; Kang et al., 2023; Wang et al., 2024).

A smaller line of work predicts adsorbate density fields, but its outputs are either conditioned on scalar uptake or normalized, and therefore do not directly determine the uptake. Sun and Siepmann (Sun & Siepmann, 2024) predict voxel-wise differential contributions to global quadratic isotherm parameters, enabling thermodynamic-state-dependent density reconstruction but requiring global uptake information. Burner et al. (Burner et al., 2025) use an equivariant graph network based on DeepDFT (Jørgensen & Bhowmik, 2020; 2022) to predict normalized adsorbate density fields in MOFs. While suitable for binding sites, this approach cannot recover global uptake N, since ρ is normalized to

integrate to unity.

3. Method

MADField is a multi-fidelity neural density operator for amortized adsorption simulation. Given a porous framework $\mathcal{M}$, an adsorbate species s, and a thermodynamic condition (P, T), MADField predicts the equilibrium adsorbate density field in physical units. Gas uptake is obtained by integrating this field, while the same prediction preserves spatial adsorption structure and can serve as an initial density for iterative cDFT solvers. The training pipeline first learns MADField-cDFT from solver-converged cDFT density fields, then uses its predictions to reinitialize failed cDFT calculations and expand the cDFT label set, and finally adapts the resulting density operator to sparse GCMC-derived density fields to obtain MADField-GCMC.

3.1. Multi-fidelity density training

Shared density target. MADField is trained to predict the equilibrium one-body density field $\rho_{\mathrm{eq}}(\mathbf{r}; \mathcal{M}, s, P, T)$. For MADField-cDFT, this density is supervised by solver-converged PC-SAFT cDFT labels; for MADField-GCMC, it is supervised by particle-to-grid averaged GCMC labels. Thus, both stages learn the same field-level object, while the source of supervision changes. The scalar uptake is recovered from the predicted density by integrating over the accessible pore region $\mathcal{P}$.

cDFT pre-training. The first stage trains MADField-cDFT on PC-SAFT cDFT density fields. These labels are available at larger scale and across more adsorbates than GCMC-derived density fields, but inherit the approximation and convergence properties of the cDFT solver. This stage learns a broad adsorption-density prior that captures how framework geometry, adsorbate identity, and thermodynamic condition shape the equilibrium density field.

cDFT refinement from predicted densities. Because MADField-cDFT predicts an unnormalized adsorbate density field, its output can also serve as an initial density for the original cDFT solver. For PC-SAFT cDFT, ρ_{eq} is the fixed point of the Picard map $\rho^{(n+1)} = \mathcal{T}_{\boldsymbol{\xi}}[\rho^{(n)}]$ with PC-SAFT parameters $\boldsymbol{\xi}$. The standard solver initializes from the Boltzmann density ρ_{B}; we instead initialize from $\rho^{(0)} = \hat{\rho}_\theta$ and keep all subsequent iterations unchanged, so refined outputs remain solver-converged cDFT densities rather than neural pseudo-labels.

Label bootstrapping. We use this refinement for label bootstrapping (Figure 2): after training MADField-cDFT on converged cDFT calculations, we reinitialize previously failed cases with its predictions and rerun the original solver, adding only newly converged densities to the supervised cDFT set.

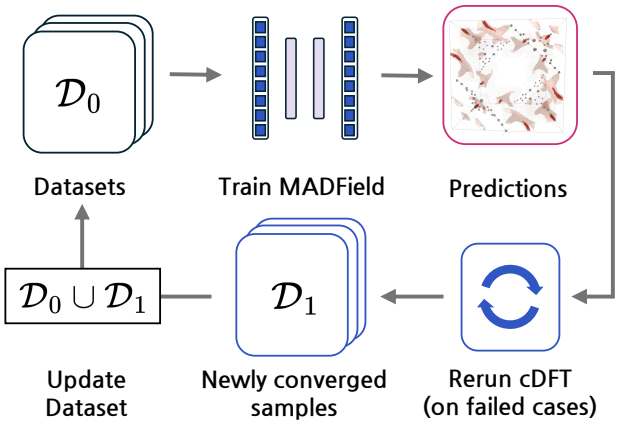

Figure 2. **Bootstrapping cDFT labels.** MADField reinitializes failed cDFT cases with predicted density fields, which are then rerun to convergence. Recovered samples are solver-converged cDFT labels, not neural pseudo-labels.

GCMC adaptation. After cDFT pre-training and bootstrapping, we adapt the resulting MADField-cDFT checkpoint to GCMC-derived density fields. GCMC provides particle-simulation references under the chosen force field, but field-level GCMC labels are more expensive and therefore sparser. MADField-GCMC is obtained by fine-tuning MADField-cDFT on paired GCMC density grids. This stage uses the cDFT-trained model as an adsorption-density prior and updates it toward particle-simulation fidelity, rather than learning the GCMC density operator from scratch. Thus, the abundant cDFT labels provide broad geometric and thermodynamic coverage, while the sparse GCMC labels correct the remaining fidelity gap.

Training objective. All training stages use the same density-field objective,

$$\mathcal{L} = \mathcal{L}_{\mathrm{field}} + \lambda_N \mathcal{L}_N, \tag{6}$$

where $\mathcal{L}_{\mathrm{field}}$ denotes the collection of voxel-wise density-supervision terms. It includes log-residual supervision relative to the Boltzmann reference, relative density error in the pore region, and auxiliary field losses that emphasize high-density adsorption sites and supervise the coarse branch against a low-pass residual target. The uptake loss $\mathcal{L}_N$ matches the integrated uptake of the prediction to that of the target density. Across cDFT and GCMC stages, the forward architecture, density parameterization, and loss form remain unchanged, only the supervision target and molecular-parameter registry differ. The full expanded loss, including all density-field terms and weights, is given in Section B. Convergence criteria and the inference-time warm-start protocol used for refinement and bootstrapping are described in Section C.

3.2. Periodic volumetric density operator

Periodic volumetric representation. The continuous density fields are represented on a periodic discrete voxel grid for learning and evaluation. The operator is designed around three requirements: periodic cell geometry, three-dimensional pore connectivity, and continuous dependence on thermodynamic and molecular conditions. The volumetric input contains the reduced external potential and the log-Boltzmann density. All boundary-facing operations are periodic: the convolutional stem uses circular padding, shifted-window attention uses circular indexing, and positional information is encoded with periodic Fourier features. We provide implementation details in Section D.

Boltzmann-residual parameterization. To encode the dominant one-body framework response, we use the ideal-gas Boltzmann density $\rho_{\text{Boltz}}(\mathbf{r}) = \rho_{\text{bulk}} \exp[-\beta V_{\text{ext}}(\mathbf{r})]$, where $\beta = (k_B T)^{-1}$. Its derivation and relation to the cDFT Euler–Lagrange equation are given in Section A. MADField predicts a gated correction in log-density space:

$$\log \hat{\rho}_\theta(\mathbf{r}) = \log \rho_{\text{Boltz}}(\mathbf{r}) + \sigma\big(h_\theta(\mathbf{r})\big)\, \delta_\theta(\mathbf{r}), \qquad (7)$$

where $\delta_\theta(\mathbf{r})$ is a residual field and $h_\theta(\mathbf{r})$ is a gate field. This guarantees $\hat{\rho}_\theta(\mathbf{r}) > 0$ and makes the network learn the many-body correction to the one-body Boltzmann response.

Backbone and conditioning. The backbone is a U-shaped 3D Swin Transformer (Liu et al., 2021). A two-stage convolutional stem downsamples the discrete voxel grid to a token resolution where windowed attention is tractable. Shifted windows, skip connections, and the decoder allow adsorption features to propagate across connected pores while keeping memory cost manageable. Non-volumetric information is injected through a conditioning vector containing $\log \rho_{\text{bulk}}$, $\log P$, lattice parameters, and a three-scalar molecular descriptor. This vector modulates every Swin block through adaptive normalization. For MADField-cDFT, the descriptor is the PC-SAFT triplet (m, σ, ε); for MADField-GCMC, the same descriptor slot is filled with Lennard-Jones parameters from the GCMC references. Full architecture details, loss weights, and parameter tables are given in Sections B, D and E.

4. Experiments

This section presents benchmark results for density field (Section 4.2) and uptake prediction (Section 4.3) in MOFs, comparing MADField with other baselines models. We further demonstrate transferability to ACs and other disordered porous materials with larger unit cells (Section 4.4), and highlight the computational speedup that MADField provides for cDFT solvers (Section 4.5).

Table 1. **Comparison of MADField with baselines.** Columns indicate support for uptake prediction (N), density-field prediction (ρ), adsorbate conditioning (Ads.), pressure conditioning (P), and material-class generalization (Mat.). ✓/✗/▲ denote full/no/partial support. MADField satisfies all five criteria.

Model	N	ρ	Ads.	P	Mat.
RetNet	✓	✗	✗	✗	✓
MOFTransformer	✓	✗	✗	✗	✗
DeepSorption	✓	✗	✗	▲	✓
IsothermNet	✓	✗	✓	✓	✓
Uni-MOF	✓	✗	✓	✓	✗
SorbIIT	✗	▲	✓	✓	✓
DeepAPD	✗	▲	✗	✗	✓
MADField (Ours)	✓	✓	✓	✓	✓

4.1. Setup

Datasets. We train and evaluate at two fidelities, cDFT and GCMC. Each fidelity provides a dataset for training and in-distribution (ID) testing, paired with a separate out-of-distribution (OOD) benchmark to assess zero-shot transferability. As part of this work, we release a large-scale cDFT adsorption dataset comprising approximately 280,000 PC-SAFT cDFT calculations across 4,000 unique MOF frameworks, nine adsorbates, and tens of pressure points per isotherm, requiring a total of 3,600 H200 GPU hours to generate. The dataset will be publicly released upon acceptance. We perform cDFT calculations to generate density fields and uptake values, using MOF structures from the QMOF database (Rosen et al., 2021) for ID (nine adsorbates) and amorphous carbon structures from Gardner et al. (Gardner et al., 2023) for OOD (CH_4). For GCMC data, we use CH_4 and Xe density fields and uptake values on a subset of the ARCMOF database from Burner et al. (Burner et al., 2022; Burner, 2025a;b;c) for ID, and CH_4 uptake on diverse disordered materials from Thyagarajan and Sholl (Thyagarajan & Sholl, 2020) for OOD, including polymers of intrinsic microporosity (PIM), hyper-cross-linked polymers (HCP), Kerogen, and ACs. Details on data generation and settings are provided in Section F.

Baselines. We compare MADField against several baseline models: RetNet (Sarikas et al., 2024), DeepSorption (Cui et al., 2023), MOFTransformer (Kang et al., 2023), Uni-MOF (Wang et al., 2024), and IsothermNet (Lin et al., 2025) for scalar uptake prediction, as well as DeepAPD (Burner et al., 2025) and SorbIIT (Sun & Siepmann, 2024) for density field prediction. Notably, DeepAPD does not support the prediction of unnormalized density fields, and SorbIIT requires gas uptake data as an input to predict them. Accordingly, these two models were excluded from the gas uptake benchmark and only assessed in the density field benchmark. For RetNet, MOFTransformer, and DeepSorption, which do not support conditioning on adsorbate and pressure, we additionally train a separate model for each (adsorbate, pressure) combination. The exception is DeepSorption, which

Table 2. **Density field prediction on MOFs.** Tanimoto similarity (↑) between predicted and ground-truth normalized density on **cDFT** densities (nine adsorbates) and **GCMC** densities (CH_4, Xe). **Total** denotes the sample-weighted mean over all evaluated density fields within each fidelity group. The conditioning columns (Cond.) indicate whether adsorbate identity (Ads.) and pressure (P) enter the model as conditioning inputs.

Cond.		Model	cDFT density										GCMC density		
Ads.	P		H_2	Ar	Kr	Xe	N_2	CH_4	CO_2	C_2H_6	C_3H_8	**Total**	CH_4	Xe	**Total**
✓	✓	SorbIIT	0.881	0.749	0.693	0.736	0.812	0.714	0.908	0.819	0.823	0.805	0.730	0.690	0.717
✗	✗	DeepAPD	0.968	0.955	0.942	0.905	0.958	0.947	0.929	0.902	0.849	0.943	0.904	0.863	0.891
✓	✓	**MADField**	**1.000**	**1.000**	**1.000**	**0.978**	**1.000**	**0.999**	**0.998**	**0.991**	**0.960**	**0.996**	**0.982**	**0.931**	**0.965**

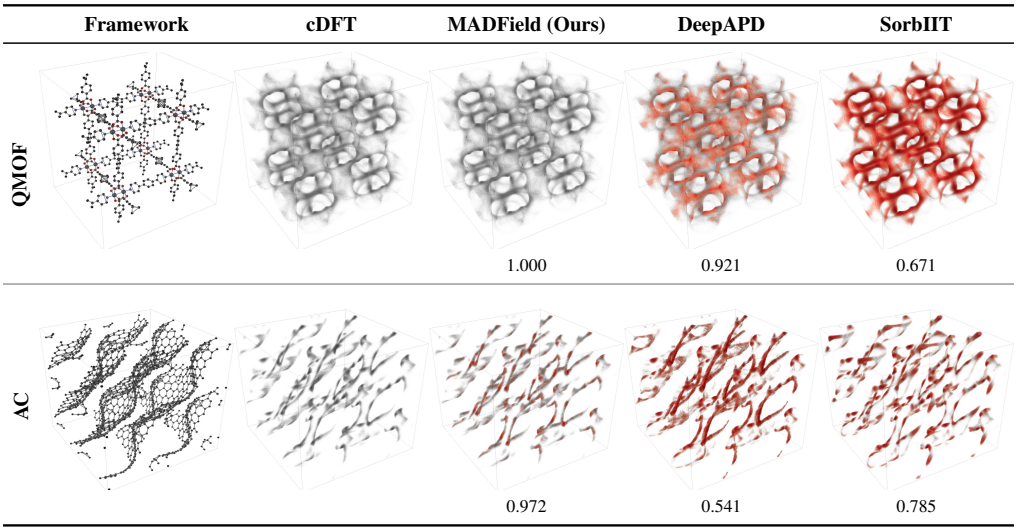

Figure 3. **Visualization of CH_4 density field and error at cDFT-level.** Rows show an ID QMOF case and an OOD AC case. Columns show the framework, **cDFT** reference density, and density errors for MADField, DeepAPD, and SorbIIT. The reference density is shown in gray, while red marks larger normalized absolute error under a shared per-row scale. Tanimoto similarity is reported below each model output. MADField remains near-white across both cases, indicating substantially lower spatial density error than the density modeling baselines.

predicts uptake at a predefined set of pressure points in a single forward pass and thus requires one model per adsorbate. Table 1 summarizes how MADField compares to existing models across five criteria: gas uptake prediction, density field prediction, adsorbate and pressure conditioning, and material-class generalization. It shows that MADField is the only model satisfying all five. Implementation details for each baseline are in Section G.

Metrics. We report gas uptake mean absolute error N_{err} in cm^3(STP)/g and Tanimoto coefficient $T_\rho \in [0, 1]$ between predicted and reference normalized densities, where $T_\rho = 0$ indicates non-overlapping densities and $T_\rho = 1$ indicates identical densities. Since DeepAPD predicts only normalized densities, we evaluate T_ρ on normalized densities for all models. Detailed metric definitions are provided in Section H.

4.2. Density fields prediction on MOFs

Table 2 reports the density prediction benchmark on the QMOF dataset (cDFT density) and a subset of ARCMOF (GCMC density). MADField consistently achieves a T_ρ close to 1 across all adsorbates, with an overall mean T_ρ of 0.996 for cDFT density and 0.965 for GCMC density. In comparison, DeepAPD follows with mean T_ρ values of 0.943 for cDFT density and 0.891 for GCMC density. SorbIIT shows lower T_ρ, falling significantly short of MADField performance. The density prediction benchmark on OOD ACs is provided in Section I, where MADField outperforms all baselines by a large margin (T_ρ: MADField 0.885 vs. SorbIIT 0.492 and DeepAPD 0.418). Figure 3 visualizes representative CH_4 density predictions on QMOF and AC. MADField produces near-zero error across both cases, while DeepAPD and SorbIIT show substantially larger errors on the OOD AC case.

Table 3. **Gas uptake prediction on MOFs.** Uptake mean absolute error N_{err} in cm^3(STP)/g (↓) on **cDFT** uptakes (nine adsorbates) and on **GCMC** uptakes (CH_4, Xe). **Total** denotes the sample-weighted mean over all evaluated state points within each fidelity group. The conditioning columns indicate whether adsorbate identity (Ads.) and pressure (P) enter as conditioning inputs.

Cond.		Model	cDFT uptake										GCMC uptake		
Ads.	P		H_2	Ar	Kr	Xe	N_2	CH_4	CO_2	C_2H_6	C_3H_8	**Total**	CH_4	Xe	**Total**
✗	✗	RetNet	4.83	3.93	9.46	22.05	3.31	17.44	17.91	27.70	46.62	12.50	10.72	10.65	10.70
✗	✗	MOFTransformer	0.12	1.87	6.21	15.09	1.32	5.32	12.12	15.66	23.40	5.76	7.22	12.32	8.92
✗	▲	DeepSorption	4.07	5.00	10.65	21.24	3.66	12.71	21.19	19.19	45.37	11.43	12.73	20.19	15.22
✓	✓	IsothermNet	5.21	7.53	18.71	40.25	5.53	21.71	44.11	52.15	81.59	21.23	16.26	21.78	18.10
✓	✓	Uni-MOF	0.46	1.81	5.11	12.21	1.22	4.13	10.39	11.81	22.12	4.93	8.11	12.29	9.50
✓	✓	**MADField**	**0.08**	**0.11**	**0.49**	**2.87**	**0.13**	**0.80**	**1.17**	**2.01**	**4.51**	**0.82**	**0.33**	**1.37**	**0.58**

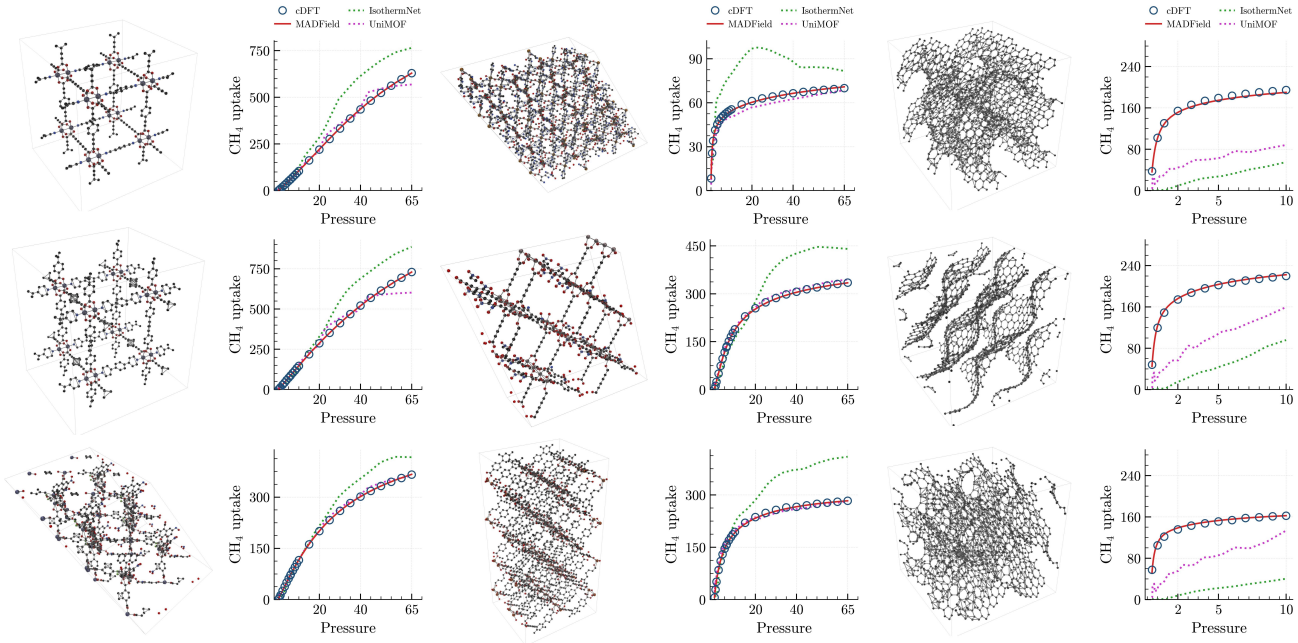

Figure 4. **Representative structures and adsorption isotherms.** Each pair of panels shows a framework structure and its corresponding CH_4 adsorption isotherm, reported as uptake in cm^3/g versus pressure in bar. The first two rows show six in-distribution QMOF examples, and the last row shows three out-of-distribution JLA amorphous-carbon examples.

4.3. Gas uptake prediction on MOFs

Table 3 reports uptake MAE on MOFs. MADField achieves $0.82\,\text{cm}^3/\text{g}$ on cDFT and $0.58\,\text{cm}^3/\text{g}$ on GCMC, improving over the strongest baselines (Uni-MOF on cDFT, MOFTransformer on GCMC) by $6.0\times$ and $15.4\times$, respectively. Figure 4 shows representative isotherm comparisons on QMOF (ID) and JLA amorphous carbons (OOD).

A backbone-shared variant MADField-Scalar that regresses uptake N directly incurs $7.2\times$ higher N_{err} than MADField, confirming that the density-field target drives most of the gain (Section I). Without cDFT pre-training, GCMC-only training reaches $1.68\,\text{cm}^3/\text{g}$ on ID MOFs ($2.9\times$ worse) and degrades to $13.60\,\text{cm}^3/\text{g}$ on OOD materials ($18.1\times$ worse than multi-fidelity), highlighting the value of the cDFT prior for generalization. On the ARCMOF GCMC subset, parity plots between predicted and reference uptake yield Pearson $r = 0.994$ (Xe, 1 bar), 0.952 (CH_4, 5.8 bar), and 0.993 (CH_4, 65 bar) (see Figure 5(a)-(c)).

4.4. Transferability to other materials

Table 4 reports zero-shot uptake MAE on disordered materials unseen during training. MADField attains $2.71\,\text{cm}^3/\text{g}$ on cDFT-AC ($4.9\times$ better than IsothermNet) and a sample-weighted total of $0.62\,\text{cm}^3/\text{g}$ on GCMC OOD ($4.21\times$ better than RetNet). These OOD materials reach up to 18,019 atoms and unit cells of 56.7 Å, yet MADField maintains sub-$1\,\text{cm}^3/\text{g}$ error across all four families. In contrast, MOF-trained baselines degrade severely (e.g., Uni-MOF rises from 7.70 to $31.30\,\text{cm}^3/\text{g}$ on cDFT-AC), whereas MAD-

Table 4. **Transferability benchmark on disordered porous materials.** Gas uptake mean absolute error N_{err} in cm^3(STP)/g (↓) on **cDFT** labels (AC at 1 bar) and on **GCMC** labels (CH_4 on PIM, HCP, Kerogen, AC). **Total** denotes the sample-weighted mean over all GCMC families. The conditioning columns indicate whether adsorbate identity (Ads.) and pressure (P) enter as conditioning inputs.

Cond.		Model	cDFT uptake	GCMC uptake			
Ads.	P		AC	PIM	HCP	Kerogen	AC	**Total**
✗	✗	RetNet	21.30	2.28	1.94	6.36	9.51	2.61
		MOFTransformer	56.40	6.90	13.10	13.60	30.60	10.03
✗	▲	DeepSorption	52.50	6.14	6.72	26.80	23.10	8.17
✓	✓	IsothermNet	13.33	9.63	11.41	44.23	46.03	13.48
		Uni-MOF	31.30	8.20	6.10	27.80	26.90	9.48
		MADField	**2.71**	**0.57**	**0.40**	**1.42**	**0.67**	**0.62**

Table 5. **MADField-initialized cDFT.** MADField increases the convergence rate and reduces Picard iterations compared with the standard Boltzmann init.

Initialization	Conv. rate (%)	Mean iter.
Standard init	90.06	329.2
MADField-cDFT	**94.21**	**166.5**

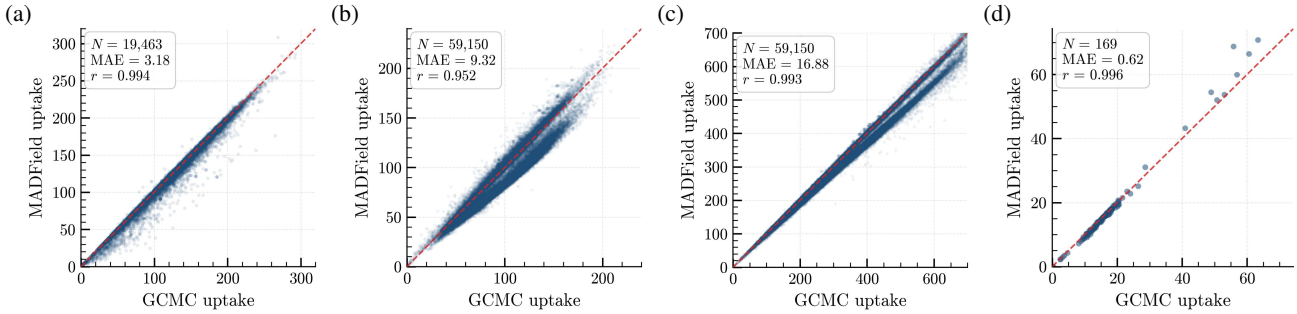

Figure 5. **Parity plots between GCMC uptake and prediction.** Each panel compares the predicted uptake against the ARCMOF **GCMC** reference which is not included in the train/val/test set: **(a)** Xe at 1 bar, **(b)** CH_4 at 5.8 bar, **(c)** CH_4 at 65 bar, and **(d)** disordered porous materials at 1 bar.

Field degrades only from 0.82 to 2.71. The OOD parity plot yields $r = 0.996$ (Figure 5(d)).

4.5. Accelerating cDFT with MADField

Accurate density predictions from MADField can serve as high-quality initial fields for iterative cDFT solvers. As shown in Table 5, MADField initialization improves the convergence rate by recovering 42% of cases that fail under standard initialization, and reduces the mean Picard iterations from 329.2 to 166.5 (2.0×). This acceleration and convergence recovery enable our cDFT-label bootstrapping strategy.

5. Conclusion

We presented MADField, a model that amortizes adsorption simulation by predicting the 3D equilibrium density field of gas molecules within nanoporous materials. By formulating gas uptake as an integral of the predicted field, MADField unifies scalar uptake prediction, spatial density estimation, and cDFT solver acceleration within a single framework. Multi-fidelity training on cDFT and GCMC data enables strong in-distribution performance on MOFs and robust generalization to out-of-distribution disordered materials, surpassing existing baselines by up to an order of magnitude in uptake accuracy.

Limitations & future work. The cDFT labels used for pre-training inherit the approximation error of the underlying free energy functional, though this is mitigated by their role as low-fidelity supervision that is subsequently corrected by GCMC fine-tuning, and by the close agreement between PC-SAFT cDFT and GCMC for the non-polar adsorbates considered here. Additionally, MADField currently treats single-component adsorption; extending to mixture predictions, for instance by combining single-component isotherms through IAST, is a promising direction for future work.

Impact Statement

This work studies learned surrogates for adsorption simulation in porous materials. By predicting equilibrium adsorbate density fields, the proposed approach may reduce the computational cost of material screening and provide spatial information that is not available from scalar property predictors. Such tools could support research on gas storage, separation, and carbon capture by helping prioritize candidates for more expensive simulations or experiments. Responsible use requires careful benchmarking, uncertainty awareness, and clear reporting of the model's domain of applicability.

References

Adams, D. Grand canonical ensemble monte carlo for a lennard-jones fluid. *Molecular Physics*, 29(1):307–311, 1975. 1, 2

Burner, J. Metal-organic framework CH4@1bar adsorbate probability distributions. https://doi.org/10.5281/zenodo.16800893, August 2025a. 5, 12, 20, 24

Burner, J. Metal-organic framework CH4@65bar adsorbate probability distributions. https://doi.org/10.5281/zenodo.16801034, August 2025b. 5, 12, 20, 24

Burner, J. Metal-organic framework Xe@1bar adsorbate probability distributions. https://doi.org/10.5281/zenodo.16801181, August 2025c. 5, 12, 20, 24

Burner, J., Luo, J., White, A., Mirmiran, A., Kwon, O., Boyd, P. G., Maley, S., Gibaldi, M., Simrod, S., and Woo, T. K. ab initio repeat charge mof database (arc-mof). https://doi.org/10.5281/zenodo.6908728, July 2022. 5, 12, 20, 24

Burner, J., Marchand, O., Cicciarella, R., Gibaldi, M., and Woo, T. K. Rapid prediction of adsorbate probability distributions in metal-organic frameworks using graph neural networks. 2025. 1, 3, 5, 12, 24

Charalambous, C., Moubarak, E., Schilling, J., Sanchez Fernandez, E., Wang, J.-Y., Herraiz, L., Mcilwaine, F., Peh, S. B., Garvin, M., Jablonka, K. M., et al. A holistic platform for accelerating sorbent-based carbon capture. *Nature*, 632(8023):89–94, 2024. 1

Chen, P., Jiao, R., Liu, J., Liu, Y., and Lu, Y. Interpretable graph transformer network for predicting adsorption isotherms of metal–organic frameworks. *Journal of Chemical Information and Modeling*, 62(22):5446–5456, 2022. 1, 3

Colón, Y. J. and Snurr, R. Q. High-throughput computational screening of metal–organic frameworks. *Chemical Society Reviews*, 43(16):5735–5749, 2014. 3

Cui, J., Wu, F., Zhang, W., Yang, L., Hu, J., Fang, Y., Ye, P., Zhang, Q., Suo, X., Mo, Y., et al. Direct prediction of gas adsorption via spatial atom interaction learning. *Nature Communications*, 14(1):7043, 2023. 3, 5

Dao, T. Flashattention-2: Faster attention with better parallelism and work partitioning. *arXiv preprint arXiv:2307.08691*, 2023. 18

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020. 18

Dubbeldam, D., Calero, S., Ellis, D. E., and Snurr, R. Q. Raspa: molecular simulation software for adsorption and diffusion in flexible nanoporous materials. *Molecular Simulation*, 42(2):81–101, 2016. 1, 12

Dufour-Décieux, V., Rehner, P., Schilling, J., Moubarak, E., Gross, J., and Bardow, A. Classical density functional theory as a fast and accurate method for adsorption property prediction of porous materials. *AIChE Journal*, 71(6):e18779, 2025. 1, 3, 14, 21

Esper, T., Bauer, G., Rehner, P., and Gross, J. Pcp-saft parameters of pure substances using large experimental databases. *Industrial & Engineering Chemistry Research*, 62(37):15300–15310, 2023. 21

Evans, R. The nature of the liquid-vapour interface and other topics in the statistical mechanics of non-uniform, classical fluids. *Advances in physics*, 28(2):143–200, 1979. 1, 3, 13

Gardner, J. L., Beaulieu, Z. F., and Deringer, V. L. Synthetic data enable experiments in atomistic machine learning. *Digital Discovery*, 2(3):651–662, 2023. 5, 20

Gross, J. and Sadowski, G. Perturbed-chain saft: An equation of state based on a perturbation theory for chain molecules. *Industrial & engineering chemistry research*, 40(4):1244–1260, 2001. 14, 21

Hjorth Larsen, A., Jørgen Mortensen, J., Blomqvist, J., Castelli, I. E., Christensen, R., Dułak, M., Friis, J., Groves, M. N., Hammer, B., Hargus, C., et al. The atomic simulation environment—a python library for working with atoms. *Journal of Physics: Condensed Matter*, 29(27):273002, 2017. 27

Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., Chen, W., et al. Lora: Low-rank adaptation of large language models. *Iclr*, 1(2):3, 2022. 16

Jørgensen, P. B. and Bhowmik, A. Deepdft: Neural message passing network for accurate charge density prediction. In *Machine Learning for Molecules Workshop@ NeurIPS 2020*, 2020. 3

Jørgensen, P. B. and Bhowmik, A. Equivariant graph neural networks for fast electron density estimation of molecules, liquids, and solids. *npj Computational Materials*, 8(1): 183, 2022. 3

Kang, Y., Park, H., Smit, B., and Kim, J. A multi-modal pre-training transformer for universal transfer learning in metal–organic frameworks. *Nature Machine Intelligence*, 5(3):309–318, 2023. 3, 5

Lee, S., Kim, B., Cho, H., Lee, H., Lee, S. Y., Cho, E. S., and Kim, J. Computational screening of trillions of metal–organic frameworks for high-performance methane storage. *ACS Applied Materials & Interfaces*, 13(20):23647–23654, 2021. 3

Lin, E., Zhong, Y., Chen, G., and Deng, S. Unified physio-thermodynamic descriptors via learned co2 adsorption properties in metal-organic frameworks. *npj Computational Materials*, 11(1):225, 2025. 3, 5

Lin, L.-C., Berger, A. H., Martin, R. L., Kim, J., Swisher, J. A., Jariwala, K., Rycroft, C. H., Bhown, A. S., Deem, M. W., Haranczyk, M., et al. In silico screening of carbon-capture materials. *Nature materials*, 11(7):633–641, 2012. 1

Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., and Guo, B. Swin transformer: Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 10012–10022, 2021. 5, 18

Loshchilov, I. and Hutter, F. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*, 2017. 16

Martin, M. G. and Siepmann, J. I. Transferable potentials for phase equilibria. 1. united-atom description of n-alkanes. *The Journal of Physical Chemistry B*, 102(14): 2569–2577, 1998. 12

Mason, J. A., Veenstra, M., and Long, J. R. Evaluating metal–organic frameworks for natural gas storage. *Chemical Science*, 5(1):32–51, 2014. 1

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32, 2019. 27

Peebles, W. and Xie, S. Scalable diffusion models with transformers. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 4195–4205, 2023. 19

Ran, Y., Sharma, S., Balestra, S., Li, Z., Calero, S., Vlugt, T., Snurr, R., and Dubbeldam, D. Raspa3: A monte carlo code for computing adsorption and diffusion in nanoporous materials and thermodynamics properties of fluids. *The Journal of Chemical Physics*, 161(11), 2024. 1, 12

Rappé, A. K., Casewit, C. J., Colwell, K. S., Goddard III, W. A., and Skiff, W. M. Uff, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *Journal of the American chemical society*, 114(25):10024–10035, 1992. 15, 21

Rehner, P., Bauer, G., and Gross, J. Feos: an open-source framework for equations of state and classical density functional theory. *Industrial & Engineering Chemistry Research*, 62(12):5347–5357, 2023. 27

Rosen, A. S., Iyer, S. M., Ray, D., Yao, Z., Aspuru-Guzik, A., Gagliardi, L., Notestein, J. M., and Snurr, R. Q. Machine learning the quantum-chemical properties of metal–organic frameworks for accelerated materials discovery. *Matter*, 4(5):1578–1597, 2021. 5, 20

Roth, R. Fundamental measure theory for hard-sphere mixtures: a review. *Journal of Physics: Condensed Matter*, 22(6):063102, 2010. 1, 3, 14

Sarikas, A. P., Gkagkas, K., and Froudakis, G. E. Gas adsorption meets deep learning: voxelizing the potential energy surface of metal-organic frameworks. *Scientific Reports*, 14(1):2242, 2024. 1, 3, 5

Sauer, E. and Gross, J. Classical density functional theory for liquid–fluid interfaces and confined systems: A functional for the perturbed-chain polar statistical associating fluid theory equation of state. *Industrial & Engineering Chemistry Research*, 56(14):4119–4135, 2017. 3, 14, 21

Shazeer, N. Glu variants improve transformer. *arXiv preprint arXiv:2002.05202*, 2020. 19

Shi, K., Li, Z., Anstine, D. M., Tang, D., Colina, C. M., Sholl, D. S., Siepmann, J. I., and Snurr, R. Q. Two-dimensional energy histograms as features for machine learning to predict adsorption in diverse nanoporous materials. *Journal of chemical theory and computation*, 19 (14):4568–4583, 2023. 3

Simon, C. M., Kim, J., Gomez-Gualdron, D. A., Camp, J. S., Chung, Y. G., Martin, R. L., Mercado, R., Deem, M. W., Gunter, D., Haranczyk, M., et al. The materials genome in action: identifying the performance limits for

methane storage. *Energy & Environmental Science*, 8(4): 1190–1199, 2015. 3

Stierle, R., Bauer, G., Thiele, N., Bursik, B., Rehner, P., and Gross, J. Classical density functional theory in three dimensions with gpu-accelerated automatic differentiation: Computational performance analysis using the example of adsorption in covalent-organic frameworks. *Chemical Engineering Science*, 298:120380, 2024. 14, 21

Sun, Y. and Siepmann, J. I. Understanding and predicting the spatially resolved adsorption properties of nanoporous materials. *Journal of Chemical Theory and Computation*, 20(12):5259–5275, 2024. 1, 3, 5, 25

Thiele, N., Teh, T. W., Bursik, B., Granderath, M., Bauer, G., Dufour-Décieux, V., Rehner, P., Stierle, R., Bardow, A., Hansen, N., et al. Efficient prediction of multicomponent adsorption isotherms and enthalpies of adsorption in mofs using classical density functional theory. *The Journal of Physical Chemistry B*, 2026. 1, 3

Thyagarajan, R. and Sholl, D. S. A database of porous rigid amorphous materials. *Chemistry of Materials*, 32(18): 8020–8033, 2020. 5, 12, 21

Wang, J., Liu, J., Wang, H., Zhou, M., Ke, G., Zhang, L., Wu, J., Gao, Z., and Lu, D. A comprehensive transformer-based approach for high-accuracy gas adsorption predictions in metal-organic frameworks. *Nature Communications*, 15(1):1904, 2024. 3, 5

Wilmer, C. E., Leaf, M., Lee, C. Y., Farha, O. K., Hauser, B. G., Hupp, J. T., and Snurr, R. Q. Large-scale screening of hypothetical metal–organic frameworks. *Nature chemistry*, 4(2):83–89, 2012. 1, 3

A. Additional Background on Adsorption Simulation

This appendix provides the theoretical and numerical background behind the density-field labels used in MADField. The main text introduces GCMC and cDFT as particle-level and density-level routes to the same equilibrium adsorption state. Here we give the corresponding ensemble definitions, density-functional equations, Boltzmann reference density, Picard iteration, PC-SAFT functional structure, and external-potential construction.

A.1. Grand Canonical Monte Carlo and Published GCMC References

Grand Canonical Monte Carlo samples adsorption equilibrium in the grand-canonical ensemble at fixed temperature T, chemical potential μ, and framework structure (Dubbeldam et al., 2016; Ran et al., 2024). For an adsorbate configuration with N molecule centers $\mathbf{r}_{1:N}$, the force-field energy can be written as

$$U_N(\mathbf{r}_{1:N}) = \sum_{i=1}^{N} V_{\text{ext}}^{\text{GCMC}}(\mathbf{r}_i) + \sum_{i<j} u_{ss}(\mathbf{r}_i, \mathbf{r}_j). \tag{8}$$

Here $V_{\text{ext}}^{\text{GCMC}}$ is the framework-fluid interaction under the GCMC force field. The term u_{ss} is the adsorbate-adsorbate interaction. A GCMC trajectory samples configurations with different particle numbers using moves such as insertion, deletion, translation, and rotation. The Markov chain is constructed so that the stationary distribution is proportional to the grand-canonical Boltzmann weight.

The equilibrium one-body density is the ensemble average of the microscopic particle density,

$$\rho_{\text{GCMC}}(\mathbf{r}) = \left\langle \sum_{i=1}^{N} \delta(\mathbf{r} - \mathbf{r}_i) \right\rangle_{T,\mu}. \tag{9}$$

In practice, this density is estimated by binning molecule centers onto a voxel grid and averaging over production samples,

$$\rho_{\text{GCMC}}(v) = \frac{1}{V_v} \langle n_v \rangle_{T,\mu}, \tag{10}$$

where v is a voxel, V_v is the voxel volume, and n_v is the number of adsorbate centers falling in that voxel.

MADField-GCMC uses published GCMC references rather than newly generated GCMC trajectories. We use ARC-MOF GCMC density fields and uptake values from Burner et al. (Burner, 2025a;b;c; Burner et al., 2022) for in-domain GCMC-fidelity training and evaluation. We use PRAM methane uptake values for zero-shot evaluation on disordered porous materials (Thyagarajan & Sholl, 2020). The force-field choices, simulation settings, and convergence protocols therefore follow the original ARC-MOF density dataset and PRAM dataset.

The MADField-GCMC molecular-conditioning slot is populated with the TraPPE-style Lennard-Jones registry(Martin & Siepmann, 1998) associated with the published GCMC references (Burner, 2025a;b;c; Burner et al., 2022). This registry is fidelity-specific and differs from the PC-SAFT registry used for MADField-cDFT. The detailed force-field parameters and simulation settings should be read from the original ARC-MOF density release and the PRAM database.

A.2. Density-Grid Convergence and Tanimoto Similarity

For density-field supervision, convergence of scalar uptake alone is not sufficient. A GCMC density field must also have a stable spatial distribution. The adsorption-probability-distribution literature uses a Tanimoto-type spatial similarity to compare two accumulated density grids (Burner et al., 2025). For two non-negative density tensors A and B defined on the same grid, the continuous Tanimoto similarity is

$$T(A, B) = \frac{\sum_i A_i B_i}{\sum_i A_i^2 + \sum_i B_i^2 - \sum_i A_i B_i + \epsilon}. \tag{11}$$

The small ϵ prevents numerical division by zero. A value of $T = 1$ indicates identical density tensors up to the grid representation. In the published GCMC density references, this type of overlap criterion is used to ensure that the spatial adsorption distribution has stabilized, not only the integrated loading.

A.3. Classical Density Functional Theory

Classical density functional theory describes adsorption equilibrium through the one-body number density field $\rho(\mathbf{r})$ (Evans, 1979). For a rigid framework, adsorbate species s, temperature T, and chemical potential μ, the grand-potential functional is

$$\Omega[\rho] = F_{\mathrm{id}}[\rho] + F_{\mathrm{exc}}[\rho; s] + \int_{\mathcal{V}} \rho(\mathbf{r}) \left[V_{\mathrm{ext}}(\mathbf{r}; s) - \mu\right] d\mathbf{r}. \tag{12}$$

Here F_{id} is the ideal-gas free-energy functional. F_{exc} is an approximate excess free-energy functional that accounts for intermolecular correlations. V_{ext} is the framework-fluid external potential. The equilibrium density is a stationary point of $\Omega[\rho]$,

$$\left.\frac{\delta\Omega[\rho]}{\delta\rho(\mathbf{r})}\right|_{\rho=\rho_{\mathrm{eq}}} = 0. \tag{13}$$

The ideal-gas functional is

$$F_{\mathrm{id}}[\rho] = k_B T \int_{\mathcal{V}} \rho(\mathbf{r}) \left[\log\left(\rho(\mathbf{r})\Lambda^3\right) - 1\right] d\mathbf{r}, \tag{14}$$

where Λ is the thermal de Broglie wavelength. Its functional derivative is

$$\frac{\delta F_{\mathrm{id}}}{\delta\rho(\mathbf{r})} = k_B T \log\left(\rho(\mathbf{r})\Lambda^3\right). \tag{15}$$

Combining Equation (13) and Equation (15) gives the Euler-Lagrange equation

$$k_B T \log\left(\rho_{\mathrm{eq}}(\mathbf{r})\Lambda^3\right) + \mu_{\mathrm{exc}}(\mathbf{r}; [\rho_{\mathrm{eq}}]) + V_{\mathrm{ext}}(\mathbf{r}) - \mu = 0, \tag{16}$$

where

$$\mu_{\mathrm{exc}}(\mathbf{r}; [\rho]) = \frac{\delta F_{\mathrm{exc}}[\rho]}{\delta\rho(\mathbf{r})} \tag{17}$$

is the local excess chemical-potential field.

The chemical potential is fixed by the bulk fluid at the same T and pressure P. Writing the bulk chemical potential as

$$\mu = k_B T \log\left(\rho_{\mathrm{bulk}}\Lambda^3\right) + \mu_{\mathrm{exc}}^{\mathrm{bulk}}, \tag{18}$$

the Euler-Lagrange equation can be rearranged into the fixed-point form

$$\rho_{\mathrm{eq}}(\mathbf{r}) = \rho_{\mathrm{bulk}} \exp\left[-\beta V_{\mathrm{ext}}(\mathbf{r}) - \beta\mu_{\mathrm{exc}}(\mathbf{r}; [\rho_{\mathrm{eq}}]) + \beta\mu_{\mathrm{exc}}^{\mathrm{bulk}}\right], \tag{19}$$

where $\beta = (k_B T)^{-1}$.

A.4. Boltzmann Reference Density

The Boltzmann reference density is the non-interacting solution obtained by dropping the excess term in Equation (19). Equivalently, it is the equilibrium density of an ideal gas in the same external potential. This gives

$$\rho_{\mathrm{Boltz}}(\mathbf{r}) = \rho_{\mathrm{bulk}} \exp\left[-\beta V_{\mathrm{ext}}(\mathbf{r})\right]. \tag{20}$$

This density captures the dominant one-body response to the framework. It is high in attractive wells, low in repulsive wall regions, and independent of fluid-fluid correlations. MADField uses this field as a physical reference and predicts a log-density residual relative to it. This makes the neural target closer to the many-body correction than to the full adsorption field.

A.5. Picard Iteration and Damped Mixing

The fixed-point equation in Equation (19) defines a nonlinear map from a current density to an updated density. For PC-SAFT parameters $\boldsymbol{\xi}$, define

$$\mathcal{T}_{\boldsymbol{\xi}}[\rho](\mathbf{r}) = \rho_{\text{bulk}} \exp\left[-\beta V_{\text{ext}}(\mathbf{r}) - \beta\mu_{\text{exc}}(\mathbf{r};[\rho]) + \beta\mu_{\text{exc}}^{\text{bulk}}\right]. \tag{21}$$

A Picard solver repeatedly evaluates this map,

$$\tilde{\rho}^{(n+1)}(\mathbf{r}) = \mathcal{T}_{\boldsymbol{\xi}}[\rho^{(n)}](\mathbf{r}). \tag{22}$$

Because the raw update can be unstable in strongly attractive or high-loading regimes, we use damped mixing,

$$\rho^{(n+1)}(\mathbf{r}) = (1-\alpha)\rho^{(n)}(\mathbf{r}) + \alpha\tilde{\rho}^{(n+1)}(\mathbf{r}), \tag{23}$$

where $\alpha \in (0, 1]$ is the damping factor. The converged solution satisfies

$$\rho_{\text{eq}} = \mathcal{T}_{\boldsymbol{\xi}}[\rho_{\text{eq}}]. \tag{24}$$

The conventional cold-start initialization is

$$\rho_{\text{cold}}^{(0)}(\mathbf{r}) = \rho_{\text{Boltz}}(\mathbf{r}). \tag{25}$$

MADField warm-starting replaces this initial density by a learned approximation to the interacting equilibrium density. The detailed warm-start schedule, clipping, and fallback protocol are given in Section C.

A.6. PC-SAFT Excess Functional

MADField-cDFT labels are generated with an inhomogeneous PC-SAFT cDFT functional (Gross & Sadowski, 2001; Sauer & Gross, 2017). The PC-SAFT excess free energy is decomposed as

$$F_{\text{exc}}[\rho] = F_{\text{hs}}[\rho] + F_{\text{chain}}[\rho] + F_{\text{disp}}[\rho] + F_{\text{assoc}}[\rho] + F_{\text{polar}}[\rho]. \tag{26}$$

The hard-sphere term F_{hs} represents excluded-volume packing and is evaluated with fundamental measure theory (Roth, 2010). The chain term F_{chain} accounts for connectivity of PC-SAFT segments. The dispersion term F_{disp} accounts for attractive fluid-fluid interactions. The association term F_{assoc} is not used because the reported MADField adsorbates are treated as non-associating fluids. The polar term F_{polar} is not used in the reported labels because we use an effective non-polar parameterization.

Each MADField adsorbate is represented by the PC-SAFT triplet

$$\boldsymbol{\xi}_{\text{cDFT}} = (m, \sigma, \epsilon). \tag{27}$$

The segment number m, segment diameter σ, and dispersion energy ϵ determine the bulk equation of state and the inhomogeneous excess functional. The same triplet is also used in the cDFT external-potential construction. The numerical parameter values are listed in Table 8.

The FMT weighted densities and PC-SAFT convolutional quantities are evaluated on the three-dimensional periodic grid. These convolutional evaluations are the main reason FFT-based implementations are natural for three-dimensional cDFT solvers (Stierle et al., 2024; Dufour-Décieux et al., 2025). This grid-based functional evaluation makes cDFT much faster than particle sampling for many adsorption state points, but the nonlinear fixed-point problem still requires iterative convergence.

A.7. cDFT External Potential

The cDFT external potential is the one-body interaction energy between an adsorbate molecule and the rigid framework. In the reported cDFT labels, it is computed from Lennard-Jones interactions between adsorbate segments and framework atoms. For a grid point $\mathbf{r}$, we use

$$V_{\text{ext}}^{\text{cDFT}}(\mathbf{r}) = m \sum_{j \in \mathcal{N}(\mathbf{r})} 4\epsilon_{fj}\left[\left(\frac{\sigma_{fj}}{r_j(\mathbf{r})}\right)^{12} - \left(\frac{\sigma_{fj}}{r_j(\mathbf{r})}\right)^{6}\right]. \tag{28}$$

Here j indexes framework atoms within the cutoff neighborhood $\mathcal{N}(\mathbf{r})$. The distance $r_j(\mathbf{r})$ is computed under periodic boundary conditions. For non-orthorhombic cells, we use fractional-coordinate wrapping and expand each structure consistently with the cutoff radius $r_c = 12.8$ Å. The fluid-framework mixed parameters are obtained with Lorentz-Berthelot mixing,

$$\sigma_{fj} = \frac{\sigma_f + \sigma_j}{2}, \qquad \epsilon_{fj} = \sqrt{\epsilon_f \epsilon_j}. \tag{29}$$

The adsorbate parameters $(m, \sigma_f, \epsilon_f)$ come from the PC-SAFT registry. The framework parameters (σ_j, ϵ_j) come from UFF (Rappé et al., 1992), with selected values listed in Table 9. The prefactor m scales the one-body framework interaction by the PC-SAFT segment count. For numerical stability, the reduced potential is clipped before being passed to the neural model and before defining pore masks.

A.8. Relation Between Theory and MADField Inputs

The theory above defines the physical inputs used by MADField. The framework geometry and adsorbate parameters determine $V_{\text{ext}}(\mathbf{r})$. The pressure and temperature determine ρ_{bulk} through the bulk equation of state. Together these define the Boltzmann reference density in Equation (20). MADField receives the reduced external potential, the log-Boltzmann density, unit-cell parameters, thermodynamic state, and molecular-conditioning vector. The model then predicts the interacting equilibrium density field by learning the residual correction that cDFT would otherwise obtain through Picard iteration or that GCMC would estimate through particle sampling.

B. Training

This appendix gives the training details for the released cDFT and GCMC models. The cDFT model is trained with supervised density labels only. The GCMC model is obtained by LoRA adaptation of the released cDFT model. Exploratory Euler–Lagrange pre-training runs were not used for the reported checkpoint and are therefore omitted.

B.1. Supervised Density Loss

Both fidelity stages use the same supervised density objective. The target density ρ^* is the solver-converged PC-SAFT cDFT density for MADField cDFT training and the GCMC density for MADField-GCMC. The pore mask is

$$\mathcal{P} = \{\mathbf{r} \mid \beta V_{\text{ext}}(\mathbf{r}) < 50\}. \tag{30}$$

The complement $\bar{\mathcal{P}}$ is treated as the wall region. The predicted and target log-residuals are

$$\Delta_\theta(\mathbf{r}) = \log \hat{\rho}_\theta(\mathbf{r}) - \log \rho_{\text{Boltz}}(\mathbf{r}), \qquad \Delta^*(\mathbf{r}) = \log \rho^*(\mathbf{r}) - \log \rho_{\text{Boltz}}(\mathbf{r}). \tag{31}$$

For numerical stability, log-domain targets are computed only inside the pore mask and are explicitly bounded. Framework voxels are excluded from log-residual losses by the pore mask $\mathcal{P} = \{\mathbf{r} \mid \beta V_{\text{ext}}(\mathbf{r}) < 50\}$. Within $\mathcal{P}$, target densities are floored before taking logarithms,

$$\log \rho^*_\epsilon(\mathbf{r}) = \log\big(\max(\rho^*(\mathbf{r}), \epsilon_{\log})\big),$$

with $\epsilon_{\log} = 10^{-12}$ for the residual loss and 10^{-30} for the hierarchical low-pass target. The resulting log-density values are hard-clamped to $[-30, 5]$. This makes the log-residual target finite for every voxel, including zero-count GCMC bins, without modifying the original density histogram used by the density and uptake losses. The log-residual target makes the residual head learn the correction on top of the Boltzmann baseline. The prediction itself still depends explicitly on ρ_{Boltz} through the output parameterization.

The full loss is

$$\mathcal{L} = \eta_{\text{res}}\mathcal{L}_{\text{res}} + \eta_\rho \mathcal{L}_\rho + \eta_{\rho,\text{top}}\mathcal{L}_{\rho,\text{top}} + \eta_N \mathcal{L}_N + \eta_{\text{hier}}\mathcal{L}_{\text{hier}}. \tag{32}$$

The residual loss is

$$\mathcal{L}_{\text{res}} = \frac{1}{|\mathcal{P}|} \sum_{\mathbf{r}\in\mathcal{P}} \big(\Delta_\theta(\mathbf{r}) - \Delta^*(\mathbf{r})\big)^2. \tag{33}$$

The pore density loss is the relative L_1 error inside the pore region,

$$\mathcal{L}_\rho = \frac{\sum_{\mathbf{r}\in\mathcal{P}} \big|\hat{\rho}_\theta(\mathbf{r}) - \rho^*(\mathbf{r})\big|}{\sum_{\mathbf{r}\in\mathcal{P}} \rho^*(\mathbf{r}) + \epsilon_\rho}. \tag{34}$$

The top-density loss applies the same relative L_1 error to the highest-density pore voxels,

$$\mathcal{L}_{\rho,\mathrm{top}} = \frac{\sum_{\mathbf{r}\in\mathcal{P}_{\mathrm{top}}} \left|\hat{\rho}_\theta(\mathbf{r}) - \rho^*(\mathbf{r})\right|}{\sum_{\mathbf{r}\in\mathcal{P}_{\mathrm{top}}} \rho^*(\mathbf{r}) + \epsilon_\rho}. \tag{35}$$

Here $\mathcal{P}_{\mathrm{top}}$ is the top 10% of pore voxels by target density. The loading loss is the absolute relative error in integrated uptake,

$$\mathcal{L}_N = \left|\frac{N[\hat{\rho}_\theta] - N[\rho^*]}{N[\rho^*] + \epsilon_N}\right|. \tag{36}$$

The hierarchical loss supervises the coarse output branch against a low-pass residual target,

$$\mathcal{L}_{\mathrm{hier}} = \frac{1}{|\mathcal{P}|}\sum_{\mathbf{r}\in\mathcal{P}} \left(c_\theta(\mathbf{r}) - \mathrm{LP}(\Delta^*)(\mathbf{r})\right)^2. \tag{37}$$

The operator LP is 3D average pooling with kernel size 4, followed by trilinear upsampling to the 128^3 grid. We use

$$(\eta_{\mathrm{res}}, \eta_\rho, \eta_{\rho,\mathrm{top}}, \eta_N, \eta_{\mathrm{hier}}) = (0.5, 2.0, 1.5, 0.25, 0.05). \tag{38}$$

The sum-rule loss used in earlier internal experiments is disabled in the released model.

B.2. MADField cDFT Training

The released cDFT model is trained in two supervised phases. The first phase trains from scratch for $50{,}000$ optimizer steps. The peak learning rate is 6×10^{-4}. The warmup length is $2{,}000$ steps. The second phase resumes from the best checkpoint of the first phase and runs for $30{,}000$ additional steps. The peak learning rate is 2×10^{-4}. The warmup length is $1{,}000$ steps. The released checkpoint is selected at step $26{,}000$ of the second phase.

Both phases use AdamW (Loshchilov & Hutter, 2017). Weight decay is 0.01, excluding bias, normalization, and positional-encoding parameters. The Adam coefficients are $\beta_1 = 0.9$ and $\beta_2 = 0.999$. The schedule is linear warmup followed by cosine decay. Training uses 8 H200 GPUs with batch size 8 per GPU. The effective batch size is 64. Gradient accumulation is 1. Gradient clipping uses max norm 1.0. Mixed precision uses bf16 autocast.

B.3. MADField-GCMC LoRA Adaptation

MADField-GCMC adapts the released cDFT model to GCMC density fidelity with LoRA (Hu et al., 2022). The base cDFT weights are frozen. LoRA adapters are inserted into `attn.qkv`, `attn.proj`, `ffn.w1`, `ffn.w2`, and `ffn.w3` in each Swin block. The LoRA rank is $r = 4$. The LoRA scale is $\alpha_{\mathrm{LoRA}} = 2r = 8$. The scaling mode is standard α_{LoRA}/r. The adapter dropout is 0.

The released MADField-GCMC run trains on paired GCMC density data for CH_4 and Xe. The training set contains $50{,}621$ paired state points. This consists of $32{,}852$ CH_4 state points and $17{,}769$ Xe state points. The validation set contains $5{,}796$ paired state points. This consists of $3{,}864$ CH_4 state points and $1{,}932$ Xe state points.

MADField-GCMC uses the same density loss as MADField cDFT training, with ρ^* replaced by the GCMC density. The molecular-conditioning vector switches from the PC-SAFT registry to the TraPPE-style registry used by the GCMC reference data. For CH_4, the MADField-GCMC conditioning triplet is $(m, \sigma, \epsilon/k_B) = (1.0, 3.73, 148.0)$. For Xe, the MADField-GCMC conditioning triplet is $(1.0, 4.10, 221.0)$. The bulk density $\rho_{\mathrm{bulk}}(P, T)$ is still computed from the PC-SAFT equation of state. This registry switch is part of the released MADField-GCMC checkpoint.

The MADField-GCMC run uses AdamW with peak learning rate 5×10^{-4}. The warmup length is $2{,}500$ steps. The total training length is $50{,}000$ steps. The best checkpoint is selected at step $45{,}000$. Training uses 8 H200 GPUs with batch size 8 per GPU. The effective batch size is 64. The released LoRA checkpoint adds $89{,}856$ adapter parameters. The total number of trainable parameters in the MADField-GCMC run is $1{,}572{,}898$ out of $6{,}443{,}410$, because output-side and conditioning modulation parameters are also trained.

C. cDFT Warm-Start Protocol

This appendix describes how MADField predictions are converted into initial conditions for the iterative cDFT solver. The cold-start baseline initializes the solver with the Boltzmann density,

$$\rho^{(0)}_{\mathrm{cold}}(\mathbf{r}) = \rho_{\mathrm{bulk}} \exp[-\beta V_{\mathrm{ext}}(\mathbf{r})]. \tag{39}$$

The warm-start run first computes the same 128^3 external-potential grid and the same PC-SAFT bulk density as the cold-start run. It then evaluates the MADField checkpoint to obtain a predicted density field $\hat{\rho}_\theta(\mathbf{r})$. Before passing this prediction to the solver, we rescale it by the pore-region mean density,

$$s = \mathrm{clip}\left(\frac{\rho_{\mathrm{bulk}}}{\frac{1}{|\mathcal{P}|}\sum_{\mathbf{r}\in\mathcal{P}} \hat{\rho}_\theta(\mathbf{r})}, 0.3, 5.0 \right). \tag{40}$$

The warm-start initial condition is then

$$\rho^{(0)}_{\mathrm{warm}}(\mathbf{r}) = \mathrm{clip}\left(s\hat{\rho}_\theta(\mathbf{r}), 10^{-30}, 0.01 \right). \tag{41}$$

The scaling prevents large global mismatch between the predicted field and the bulk density from destabilizing the first Picard steps. The upper density clip is used only for numerical initialization stability.

Both cold and warm runs solve the same PC-SAFT cDFT fixed-point equation. They differ only in the initial density and the Picard damping schedule. The cold-start schedule is conservative,

$$(0.005, 500), \quad (0.01, 2000), \quad (0.05, 5000), \tag{42}$$

where each pair denotes damping factor and maximum iteration count. The warm-start schedule uses larger damping factors and a lower iteration cap,

$$(0.01, 200), \quad (0.05, 1000), \quad (0.10, 3000). \tag{43}$$

The cold-start cap is 7500 iterations. The warm-start cap is 4200 iterations. Both use the same convergence tolerance of 10^{-5}.

If the warm-start run diverges or collapses to an unphysical low-density solution, the solver falls back to the cold-start schedule. This fallback is used only to make the warm-start protocol robust. The iteration-count speedups reported in the main text are computed on state points that converge under both initializations. Thus the reported speedup measures the reduction in Picard iterations when the same cDFT functional is solved from a better initial density field.

For multi-pressure isotherms, pressure continuation is separate from the MADField initialization. When a lower-pressure cDFT solution is already available, it can initialize the next pressure step through a pressure ladder. The MADField warm-start analysis in the main text isolates the effect of replacing the Boltzmann initialization by the learned density field.

D. Architecture Details

This appendix provides implementation details for the MADField volumetric density operator. The main text describes the architecture at the level needed to understand the method. Here we report the input construction, tokenization, periodic attention, conditioning, decoder, output head, and parameter counts used in the reported model.

D.1. Input Channels

MADField takes a 128^3 periodic voxel grid with two volumetric input channels. The first channel is the reduced external potential,

$$\tilde{V}(\mathbf{r}) = \frac{\mathrm{clamp}\big(\beta V_{\mathrm{ext}}(\mathbf{r}), -15, 60\big)}{15}. \tag{44}$$

The second channel is the clipped log-Boltzmann density,

$$\tilde{\rho}_B(\mathbf{r}) = \frac{\mathrm{clamp}\big(\log \rho_{\mathrm{bulk}} - \beta V_{\mathrm{ext}}(\mathbf{r}), -30, 5\big)}{15}. \tag{45}$$

The upper clamp on βV_{ext} makes strongly repulsive wall voxels saturate to the same value, which gives the network an implicit wall signal without adding a separate binary mask. The pore mask threshold $\beta V_{\text{ext}} < 50$ is used only for loss and evaluation and is described in Section B. Although the two channels are analytically related before clipping, they play different numerical roles after normalization: βV_{ext} provides a pressure-independent wall and well geometry, while $\log \rho_B$ provides the pressure-dependent ideal-gas density scale used by the residual reconstruction.

D.2. Fixed-Grid Representation

All reported MADField models represent each periodic simulation cell on a 128^3 voxel grid. This gives a common tensor interface for batching heterogeneous unit cells and matches the grid-based cDFT and processed GCMC density labels used in the paper. For a cell with maximum side length $L_{\max}$, the largest voxel spacing is approximately $L_{\max}/128$. In our datasets, even large cells with $L_{\max} \lesssim 60$ Å therefore have spacing below 0.47 Å, which is comparable to or finer than the 0.5 Å-scale resolution commonly used for adsorption energy or density grids. For the QMOF structures used in the main cDFT benchmark, cell dimensions are typically below this upper bound, giving finer effective spacings than the worst-case 60 Å cell. The lattice lengths and angles are also provided in the non-volumetric conditioning vector, allowing the model to condition on the physical scale and shape represented by the fixed voxel grid.

D.3. ConvStem

A two-stage 3D convolutional stem tokenizes the 128^3 grid into a 32^3 token grid. Both convolutions use circular padding to respect periodic boundary conditions. The stem is

$$\begin{aligned}&\text{Conv3d}(2 \to 48,\ k=3,\ s=2) \to \text{GroupNorm}(1, 48) \to \text{GELU},\\ &\text{Conv3d}(48 \to 288,\ k=3,\ s=2) \to \text{GroupNorm}(1, 288) \to \text{GELU}.\end{aligned} \tag{46}$$

The resulting token grid has $32^3 = 32{,}768$ tokens. Each token corresponds to a 4^3 voxel patch in the original density grid, following the patch-tokenization paradigm used in vision transformers (Dosovitskiy et al., 2020).

D.4. Periodic Fourier Positional Encoding

We use a periodic Fourier positional encoding on the 32^3 token grid. The encoding satisfies $\text{PE}(\mathbf{x}) = \text{PE}(\mathbf{x} + \mathbf{L})$ by construction. For a token index p along one axis with $n = 32$ tokens, integer harmonics are encoded as

$$\text{PE}_{\text{axis}}[p, 2k] = \sin\left(\frac{2\pi h_k p}{n}\right) \text{decay}_k, \qquad \text{PE}_{\text{axis}}[p, 2k+1] = \cos\left(\frac{2\pi h_k p}{n}\right) \text{decay}_k. \tag{47}$$

Integer frequencies guarantee exact periodicity across opposite faces of the unit cell. The three axis-wise encodings are concatenated to match $d_{\text{model}} = 288$.

D.5. Conditioning Vector

The non-volumetric conditioning vector $\mathbf{c} \in \mathbb{R}^{11}$ concatenates the bulk thermodynamic state, the unit-cell axes, the unit-cell angle deviations from $90°$, and the adsorbate triplet,

$$\mathbf{c} = \text{standardize}\big(\rho'_{\text{bulk}},\ P',\ a',\ b',\ c',\ \alpha',\ \beta',\ \gamma',\ m',\ \sigma',\ \varepsilon'\big),$$

where $x' = \log x$ for the eight scale variables $\{\rho_{\text{bulk}}, P, a, b, c, m, \sigma, \varepsilon\}$ and $\theta' = \theta - 90°$ for the three cell angles $\{\alpha, \beta, \gamma\}$. Standardization statistics are computed once on the training split and stored as fixed buffers. The angle-deviation parameterization places an orthorhombic cell at the origin of the angular axes. For cDFT training, (m, σ, ε) is the PC-SAFT parameter triplet. For GCMC fine-tuning, the same three slots are populated by the TraPPE-style registry of Section A.

D.6. Periodic Swin Blocks

The backbone contains $L = 6$ Swin-style blocks with $d_{\text{model}} = 288$ (Liu et al., 2021). Each attention layer uses 8 heads, giving head dimension 36. Window attention uses windows of size 8^3 on the 32^3 token grid. This gives $4^3 = 64$ windows with 512 tokens per window. Attention is evaluated by `scaled_dot_product_attention`, using the FlashAttention-2 backend on H100 and H200 GPUs when available (Dao, 2023). Odd-indexed blocks shift the token grid by $w/2 = 4$ tokens

along all three axes using `torch.roll`. Because the grid is periodic, rolled tokens are physically adjacent across unit-cell boundaries. No attention mask is used.

Each block uses a SwiGLU feed-forward network (Shazeer, 2020). The nominal FFN width is $d_{\text{ff}} = 2d_{\text{model}} = 576$. The SwiGLU inner width is

$$d_{\text{inner}} = \text{round\_up\_to\_8}\left(\frac{2d_{\text{ff}}}{3}\right) = 384. \tag{48}$$

This keeps the SwiGLU parameter count comparable to a standard FFN with expansion ratio 2.

D.7. AdaLN-Zero Conditioning

Each Swin block is modulated by AdaLN-Zero (Peebles & Xie, 2023). The conditioning vector $\mathbf{c} \in \mathbb{R}^{11}$ is mapped by a three-layer MLP,

$$11 \to 128 \to 128 \to 6d_{\text{model}}. \tag{49}$$

The output dimension is $6d_{\text{model}} = 1728$. It provides $(\gamma_1, \beta_1, \alpha_1)$ for attention and $(\gamma_2, \beta_2, \alpha_2)$ for the FFN. For the attention sublayer, the update is

$$y = (1 + \gamma_1) \odot \text{LN}(x) + \beta_1, \qquad x \leftarrow x + \alpha_1 \odot \text{Attn}(y). \tag{50}$$

The FFN sublayer uses the same form with $(\gamma_2, \beta_2, \alpha_2)$. The final linear layer of the AdaLN MLP is zero-initialized. Thus γ, β, and α start at zero, and each block is initialized as an identity residual block.

D.8. U-Shape Decoder

The six Swin blocks are arranged as a shallow U-shape. Blocks $0, 1, 2$ form the encoder. Blocks $3, 4, 5$ form the decoder. Encoder hidden states are cached and concatenated with the corresponding decoder states. Each skip connection uses a linear projection

$$\text{Linear}(2d_{\text{model}}, d_{\text{model}}). \tag{51}$$

There are three such projections. Together they contain $498{,}528$ parameters.

D.9. Output Head

The output head reconstructs a 128^3 log-density residual from the final 32^3 token grid. It contains a fine residual head, a gate head, and a coarse low-frequency head. The full output equation is

$$\log \hat{\rho}_\theta(\mathbf{r}) = \log \rho_{\text{Boltz}}(\mathbf{r}) + \alpha\, \sigma\big(g_\theta(\mathbf{r})\big) \left[r_\theta(\mathbf{r}) + \lambda\, \text{Up}\left(r_\theta^{\text{coarse}}\right)(\mathbf{r})\right]. \tag{52}$$

Here $\alpha = 1.0$ is a fixed registered buffer and is not learned. The fine residual head is $\text{Linear}(d_{\text{model}}, 64)$, producing one 4^3 patch per token. The gate head is $\text{Linear}(d_{\text{model}}, 64)$, also producing one 4^3 patch per token. The coarse head is $\text{Linear}(d_{\text{model}}, 1)$, producing one scalar per token on the 32^3 token grid. The coarse output is trilinearly upsampled to the 128^3 voxel grid before being added to the fine residual. The scalar λ is an unconstrained learned parameter initialized to 0.3. It is not passed through a sigmoid and is not constrained to $[0, 1]$.

The residual head uses Gaussian initialization with standard deviation 0.02. The gate head has zero weights and bias -2, so the initial gate value is $\sigma(-2) \approx 0.12$. The coarse head is zero-initialized. This makes the initial prediction close to the Boltzmann baseline while still allowing a small learned correction from the residual branch.

After reconstruction, the log-density is clamped per sample to

$$\log \hat{\rho}_\theta \in [\log \rho_{\text{bulk}} - 20, \log \rho_{\text{bulk}} + 10]. \tag{53}$$

This numerical clamp prevents extreme densities while preserving the orders-of-magnitude variation across pressure conditions.

D.10. Parameter Summary

The final model has 6.35M parameters. Most parameters are in the Swin blocks. The approximate module-level breakdown is shown in Table 6, and model hyperparameters are shown in Table 7.

Table 6. **Approximate parameter breakdown of MADField.**

Module	Share of parameters
Swin blocks	86%
U-shape skip projections	8%
ConvStem and input processing	6%
Output heads	$< 1\%$
Coarse blend parameter λ	1 parameter

Table 7. **MADField architecture hyperparameters.**

Component	Value
Input grid	128^3
Token grid	32^3
Patch size	4^3 voxels
d_{model}	288
Number of Swin blocks	6
Encoder blocks	3
Decoder blocks	3
Attention heads	8
Head dimension	36
Window size	8^3 tokens
FFN type	SwiGLU
Nominal FFN width	576
SwiGLU inner width	384
Conditioning dimension	11
AdaLN MLP	$11 \to 128 \to 128 \to 1728$
Drop-path rate	0.1 linear schedule
Residual head	$\text{Linear}(288, 64)$
Gate head	$\text{Linear}(288, 64)$
Coarse head	$\text{Linear}(288, 1)$
Gate bias	-2
Output α	fixed 1.0
Coarse blend λ	raw learned scalar initialized to 0.3
Output clamp	$[\log \rho_{\text{bulk}} - 20, \log \rho_{\text{bulk}} + 10]$
Total parameters	6.35M

E. PC-SAFT Parameters and Selected Constants

We show our chosen PC-SAFT parameters and selected constants here.

F. Data Generation Pipeline

This appendix summarizes the data sources and label-generation pipeline used for the cDFT and GCMC experiments. MADField-cDFT uses cDFT density fields generated by our PC-SAFT cDFT solver. MADField-GCMC uses published GCMC references from ARC-MOF and PRAM. All density fields are represented on 128^3 periodic voxel grids unless otherwise stated.

F.1. Dataset Sources

We use four material sources in the paper. QMOF provides the main crystalline-MOF dataset for cDFT training and in-domain evaluation (Rosen et al., 2021). ARC-MOF provides published GCMC density and uptake references for the GCMC-fidelity benchmark (Burner et al., 2022; Burner, 2025a;b;c). The JLA-Gardner amorphous-carbon ensemble provides an out-of-domain cDFT transfer set (Gardner et al., 2023). PRAM provides out-of-domain GCMC uptake references for

Table 8. PC-SAFT non-polar parameters for the nine adsorbates. All species use the Esper non-polar set (Esper et al., 2023).

Adsorbate	m	σ (Å)	ε/k_B (K)	M_w (g/mol)
H_2	1.000	2.960	34.20	2.016
Ar	1.000	3.378	117.81	39.95
Kr	1.000	3.608	164.02	83.80
Xe	1.000	3.927	227.70	131.29
CH_4	1.000	3.701	150.07	16.04
N_2	1.238	3.300	89.41	28.01
C_2H_6	1.607	3.517	191.45	30.07
C_3H_8	1.986	3.624	209.09	44.10
CO_2	2.531	2.579	153.32	44.01

Table 9. Selected Universal Force Field (Rappé et al., 1992) Lennard-Jones parameters for framework atoms. The full UFF table covers all 102 elements; values for non-listed elements follow directly from Rappé et al. (1992).

Element	σ (Å)	ε/k_B (K)	Element	σ (Å)	ε/k_B (K)
C	3.431	52.84	O	3.118	30.19
H	2.571	22.14	N	3.261	34.72
S	3.595	137.88	F	2.997	25.16
Cl	3.516	114.23	Si	3.826	202.29
Zn	2.462	62.40	Cu	3.114	2.52
Zr	2.783	34.72	Fe	2.594	6.54
Co	2.559	7.05	Ni	2.525	7.55
Al	4.008	254.13	Mg	2.691	55.86

disordered nanoporous materials (Thyagarajan & Sholl, 2020).

QMOF. We use the QMOF database, whose structures are provided as DFT-relaxed MOF geometries. After structure parsing, pore filtering, and simulation filtering, the cDFT dataset contains 5,770 unique MOFs. This prevents leakage across different adsorbates or pressures of the same material.

ARC-MOF. The GCMC in-domain benchmark uses ARC-MOF structures with published GCMC adsorption references. For the present benchmark, we use the available GCMC density and uptake data for CH_4 and Xe. These labels define the target fidelity for MADField-GCMC.

JLA-Gardner amorphous carbon. The cDFT out-of-domain transfer benchmark uses amorphous carbon structures from the JLA-Gardner ensemble. These structures are non-crystalline carbon networks and do not share the metal-linker motifs present in QMOF. We generate cDFT references for this set using the same cDFT pipeline as QMOF.

PRAM. The GCMC out-of-domain transfer benchmark uses the PRAM disordered nanoporous dataset. We use PRAM methane uptake references across PIMs, HCPs, kerogens, and amorphous carbons. PRAM is used only for uptake evaluation in the main paper, not for density-field supervision.

F.2. cDFT Reference Generation

MADField-cDFT data labels are generated by solving PC-SAFT cDFT on a periodic 128^3 grid (Gross & Sadowski, 2001; Sauer & Gross, 2017; Stierle et al., 2024; Dufour-Décieux et al., 2025). The solver uses a non-polar PC-SAFT decomposition with FMT hard-sphere, hard-chain, and dispersion contributions.

Each calculation starts from the Boltzmann density

$$\rho_{\mathrm{Boltz}}(\mathbf{r}) = \rho_{\mathrm{bulk}} \exp[-\beta V_{\mathrm{ext}}(\mathbf{r})]. \tag{54}$$

The Picard iteration then solves the cDFT fixed-point equation until convergence or until the iteration cap is reached. Pressure points within the same isotherm are evaluated from low to high pressure, using the nearest lower-pressure converged solution as a warm start when available. Failed pressure points are excluded from supervised training but retained for the warm-start recovery analysis in Section C.

Table 10. **cDFT pressure grids at $T = 298.15$ K.**

Fluid	# pressures	Pressure range
Ar, C_2H_6, C_3H_8, CO_2, Kr, N_2, Xe	12	$0.1, 0.5, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10$ bar
CH_4	24	$0.1, 0.5, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10,$ $15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65$ bar
H_2	30	$0.1, 0.5, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10,$ $15, 20, 25, 30, 35, 40, 45, 50,$ $55, 60, 65, 70, 75, 80, 85, 90, 95, 100$ bar

External potential. For each framework and adsorbate pair, we precomputed one external-potential grid and reuse it across all pressures. The potential is a Lennard-Jones 12-6 framework-fluid interaction with Lorentz-Berthelot mixing. Framework atoms use UFF parameters. Adsorbates use the cDFT PC-SAFT parameter registry described in Section A. The cutoff is $r_c = 12.8$ Å. For triclinic and non-orthorhombic cells, the simulation cell is expanded consistently with this cutoff and periodic distances are computed under the minimum-image convention. The resulting V_{ext} is stored in Kelvin, and βV_{ext} is stored as a normalized model input.

Pressure grids. All cDFT data are generated at $T = 298.15$ K, and the base grid is shown in Table 10.

F.3. GCMC Reference Data

We use the released GCMC density from ARC-MOF and released GCMC uptake from PRAM in MADField-GCMC. The ARC-MOF references provide density grids and uptake values for in-domain GCMC-fidelity training and evaluation. The PRAM references provide methane uptake values for out-of-domain evaluation on disordered nanoporous materials.

ARC-MOF GCMC references. The in-domain GCMC benchmark uses ARC-MOF structures with published GCMC density and uptake references. We use the available CH_4 and Xe GCMC data for MADField-GCMC training and evaluation. The density fields are mapped to the same 128^3 grid convention used by the model. The corresponding uptake values are used for scalar GCMC benchmark comparisons.

PRAM GCMC references. The PRAM benchmark is used only for zero-shot uptake evaluation. It contains methane GCMC uptake references across PIMs, HCPs, kerogens, and amorphous carbons. No PRAM structures are used for MADField-GCMC training. Because PRAM methane loadings are smaller than the high-pressure ARC-MOF MOF loadings, we report both absolute MAE and mean relative error in the main text.

G. Baseline Implementations

This appendix describes how scalar and density adsorption baselines are adapted to our cDFT and GCMC benchmarks. All baselines are trained and evaluated on the same train, validation, and test splits as MADField whenever the model interface allows it. We keep each baseline architecture unchanged unless an implementation fix is required for training or inference. Because the original scalar models differ in whether they accept adsorbate and pressure as inputs, we use the most faithful task formulation supported by each model.

G.1. Scalar Uptake Baselines

RetNet. RetNet is used as a voxel-CNN scalar baseline. We keep the original architecture and training recipe unchanged. The model consists of five 3D convolutional layers, two max-pooling layers, and three fully connected layers, with BatchNorm and LeakyReLU activations. The first convolution uses circular padding to respect periodic boundary conditions. The model has 0.367M parameters.

RetNet does not condition on adsorbate identity or pressure. We therefore train a separate RetNet model for each $(\text{adsorbate}, P)$ condition. For the cDFT benchmark, this produces one model per gas and pressure state. For the GCMC benchmark, we train separate models for the available CH_4 and Xe conditions. The input voxel transform follows the

original RetNet setup:

$$X = \mathrm{clip}(-U/T_{\mathrm{room}}, 0, 50),$$

followed by standardization using the training-set mean and standard deviation. We use the original random rotation, flip, reflection, and identity augmentations. We also keep the original optimizer and schedule, using Adam with learning rate 10^{-3}, StepLR with step size 10 and decay factor 0.5, MSE loss, and 30 training epochs. The final-layer bias is initialized to the training-label mean, as in the original implementation. Labels are trained in raw uptake units without target normalization. We store the training-set voxel normalization statistics with each checkpoint and use them at evaluation time.

DeepSorption. DeepSorption is used through its Matformer implementation. We keep the published architecture unchanged, including the $N = 8$ encoder layers, embedding dimension 512, feed-forward dimension 2048, 4 attention heads, constrained attention distance bins, pre-trained RotatE embeddings, 3D positional encoding, and generator head. The resulting model has 26.06M parameters.

DeepSorption predicts a vector of scalar targets for a given structure. We therefore train one DeepSorption model per adsorbate, with the output dimension set to the number of pressure points for that adsorbate. This allows the model to predict a pressure grid for a fixed gas, while avoiding an unsupported pressure-conditioned regression interface. For cDFT, the output dimension varies across gases because the pressure grid differs by adsorbate. For GCMC, the output dimension matches the available pressure points in the published GCMC benchmark.

The main changes are limited to the data and training pipeline. Targets are pivoted from the long-format $(\mathrm{MOF}, \mathrm{adsorbate}, P)$ table into per-gas pressure vectors. Missing or unconverged cDFT state points are represented by a binary mask and excluded from the loss. The training loss is masked MSE,

$$\mathcal{L}_{\mathrm{DS}} = \frac{\sum_k m_k \left(\hat{y}_k - y_k\right)^2}{\sum_k m_k + \epsilon},$$

where $m_k = 1$ for valid pressure points and $m_k = 0$ for missing or unconverged labels. Target normalization follows the original per-target z-score strategy, but statistics are computed only over valid training labels. We use the same ReduceLROnPlateau scheduler as the original code, with factor 0.6, patience 10, and minimum learning rate 10^{-7}. The batch size is reduced to 16 because the maximum sequence length is increased for our MOF structures. We select checkpoints using a robust validation uptake error computed only on state points with reference loading above 0.05 mol kg^{-1}, which avoids unstable relative errors in the near-zero loading regime. We remove the evaluation-time ReLU applied to the first output column in the original script and use the raw regression outputs for all pressure points.

MOFTransformer. We use the pip-installable `moftransformer==2.1.4` package without modifying the source code. The transformer backbone, data modules, featurization utilities, and regression head are kept unchanged. We add wrapper scripts only for data preparation, launching fine-tuning runs, and saving per-structure predictions.

MOFTransformer does not take pressure, temperature, or adsorbate identity as regression inputs in its scalar head. We therefore fine-tune a separate checkpoint for each $(\mathrm{adsorbate}, P)$ condition. Each checkpoint starts from the same pre-trained MOFTransformer backbone and predicts one scalar uptake target per MOF. The target is normalized with the z-score statistics computed from the corresponding training split. For cDFT, we train separate checkpoints for the gas and pressure conditions evaluated in the benchmark. For GCMC, we train separate checkpoints for the available CH_4 and Xe conditions.

Input preparation follows the upstream MOFTransformer pipeline. Each structure is converted into the CGCNN graph representation and GRIDAY Lennard-Jones probe grid used by the original package. We parallelize the upstream data-preparation routine across CPU workers, but do not change the generated features. For test evaluation, we wrap the upstream test-only path to save per-MOF predictions to CSV, since the original script logs aggregate metrics but does not export all predictions. For amorphous-carbon OOD inference, we use an inference-only CIF parsing bypass to avoid failures on large P1 cells. This bypass is not used for the crystalline QMOF or ARC-MOF training data.

Uni-MOF. We use the released Uni-MOF code with minimal task-registration changes. The transformer backbone, UniMat atom-pair encoder, data loader, loss function, and target normalization scheme are kept unchanged. We only register new task identifiers for our cDFT and GCMC datasets by adding target normalization statistics and pressure-temperature normalization ranges to the existing registry dictionaries.

Unlike MOFTransformer and RetNet, Uni-MOF natively conditions on adsorbate and thermodynamic variables. Its input includes a gas identifier, gas attribute vector, $\log_{10} P$, and temperature. We therefore train one checkpoint for the multi-gas cDFT benchmark and one checkpoint for each GCMC training setup. The training LMDB stores atoms, coordinates, lattice matrix, gas identifier, gas attributes, temperature, pressure, target uptake, and task name for each state point. Target values use the upstream log1p-standardization transform. At inference time, the upstream inverse transform is used to recover uptake in physical units.

We add wrapper scripts to build LMDB files from our CSV splits, compute target statistics, register tasks idempotently, launch distributed training, and merge shard-level inference outputs. No architectural or algorithmic change is made to the released Uni-MOF model.

IsothermNet. We adapt the published IsothermNet implementation for multi-gas adsorption prediction. The original implementation is designed around a single-gas CO_2 isotherm setting with hard-coded gas constants in the structural global features. We replace these constants with per-sample molecular parameters so that the same architecture can receive the appropriate gas information for each state point. For our cDFT benchmark, the molecular slots are populated with the nine-gas PC-SAFT registry used elsewhere in the paper. For the GCMC benchmark, the slots are populated with the corresponding GCMC molecular registry.

We make several implementation-level fixes needed for stable training. Modules that were instantiated inside the forward pass are moved into `__init__` so that their parameters and batch-normalization statistics are registered and updated during training. We also use float32 training, vectorize repeated tensor operations in the forward pass, and reconstruct batch indices inside the model for reliable multi-GPU execution. These changes do not alter the intended model architecture but make training deterministic and reproducible.

The original feature-generation pipeline relies on molecular files and RDKit-based bond perception, which is not robust for our MOF, amorphous-carbon, and PRAM structures. We therefore rewrite the featurization pipeline using pymatgen structures, periodic distance matrices, covalent-radius bond detection, graph-based ring detection, and parallel structure processing. The Zeo++ textural feature pipeline is also automated to produce the textural inputs required by IsothermNet. We train IsothermNet with log-space MSE on $\log(q + 10^{-3})$, use per-gas and per-pressure validation metrics, and convert predictions back to $\mathrm{cm}^3(\mathrm{STP})\ \mathrm{g}^{-1}$ for the benchmark tables.

Summary. The scalar baselines differ in the conditioning axes they support. RetNet and MOFTransformer do not natively condition on pressure and adsorbate identity, so we train separate checkpoints for unsupported axes. DeepSorption supports multi-pressure output for a fixed gas, so we train one checkpoint per adsorbate. Uni-MOF and IsothermNet support molecular and thermodynamic conditioning after task registration or molecular-parameter injection, so they are trained as conditional scalar predictors. These choices give each scalar baseline the strongest task formulation supported by its released architecture while preserving a fair comparison to MADField.

G.2. Density-Field Baselines

DeepAPD. DeepAPD predicts normalized adsorbate probability distributions for binding-site localization in MOFs (Burner et al., 2025). We retain the released PaiNN-Charge backbone, including the atom and probe encoders, hidden dimension, radial basis cutoff, Softplus output, and disconnected-probe masking policy. We adapt only the training target, loss, and inference normalization so that the model can be evaluated on our physically scaled density benchmarks.

The original DeepAPD objective normalizes both prediction and target before computing a Tanimoto-style loss. This is appropriate for adsorption probability distributions, but removes the physical density scale needed for uptake integration. We therefore remove target normalization in the LMDB construction and train on physically scaled density,

$$\tilde{\rho}(\mathbf{r}) = \rho(\mathbf{r})/\rho_{\mathrm{norm}}, \qquad \rho_{\mathrm{norm}} = 5 \times 10^4\ \mathrm{mol\,m^{-3}}.$$

The predicted field is multiplied by ρ_{norm} to recover density in physical units.

For GCMC evaluation, we use public checkpoint where DeepAPD is trained on the same published Burner et al. ARC-MOF density references used by MADField-GCMC (Burner et al., 2022; Burner, 2025a;b;c). The released DeepAPD checkpoint contains task heads for CH_4 at 1 bar, CH_4 at 65 bar, and Xe at 1 bar. We use the heads matching our GCMC benchmark and retrain the probe heads for our supervised target. This gives a field directly comparable to MADField outputs.

SorbIIT. SorbIIT predicts spatially resolved adsorption information using a periodic-aware 3D U-Net architecture (Sun & Siepmann, 2024). We use the released `CellUNet` architecture with periodic unit-cell convolutions and the same channel configuration. Since the released code does not provide pre-trained checkpoints or public training data for our benchmark setting, we retrain SorbIIT from scratch on our splits.

The original SorbIIT pipeline takes proprietary zeolite energy-grid inputs and predicts voxel-wise isotherm coefficients. We replace the ZTB energy grid with our cDFT external-potential grid V_{ext}/k_B. Following the released preprocessing convention, we clamp the input to $[-10^4, 10^4]$ K and normalize it as

$$x(\mathbf{r}) = -\frac{V_{\mathrm{ext}}(\mathbf{r})/k_B - 6400}{6400}.$$

For the cDFT benchmark, the volumetric fields are resampled to a fixed physical resolution following the SorbIIT convention, with density grids rescaled to preserve the cell integral.

Because our benchmark is fixed-temperature adsorption, we use the fixed-temperature quadratic isotherm parameterization. The network predicts voxel-wise isotherm coefficient fields, from which density fields are reconstructed at the target pressures. We train with a hybrid objective combining coefficient reconstruction, direct density supervision, integrated-loading supervision, non-negativity regularization, and the SorbIIT amplitude regularizer. In our main cDFT runs, the coefficient loss is down-weighted and direct density supervision is dominant ($\lambda_X = 0.1$, $\lambda_\rho = 1.0$, $\lambda_N = 0.1$, $\lambda_{\mathrm{neg}} = 0.05$, and regularization weight 0.05).

For cDFT comparison, we train one SorbIIT model per fluid on the same train, validation, and test split used by MADField-cDFT. For GCMC comparison, we train on the same ARC-MOF density references used by the other GCMC-fidelity models. For amorphous-carbon evaluation, crystallographic symmetry augmentation is disabled and only the identity transform is used, since these cells do not have the crystallographic symmetry operations assumed in the original zeolite setting.

Faithfulness of density baselines. Both density baselines preserve their released model architectures. DeepAPD keeps the PaiNN-Charge architecture, and SorbIIT keeps the `CellUNet` architecture. The modifications are restricted to data formatting, target scaling, loss definitions, and inference reconstruction so that both models can be evaluated on the same physically scaled density task as MADField. s

H. Evaluation Metrics

We use two primary evaluation metrics. Uptake accuracy is measured by mean absolute error in gas loading. Density-field quality is measured by Tanimoto similarity after normalizing each density field to unit mass.

Uptake MAE. For each state point i, the reference loading is q_i and the predicted loading is $\hat{q}_i$, both reported in cm^3(STP) g^{-1}. For density-prediction models, $\hat{q}_i$ is obtained by integrating the predicted density field over the accessible pore region and converting the result using the sample-specific unit-cell volume and framework mass. For scalar baselines, $\hat{q}_i$ is the directly predicted uptake after unit conversion when needed. We report

$$\mathrm{MAE}_q = \frac{1}{N}\sum_{i=1}^{N} |\hat{q}_i - q_i| \,. \tag{55}$$

Lower values indicate more accurate uptake prediction.

Density Tanimoto similarity. We evaluate spatial density-field quality using the continuous Tanimoto coefficient. For each sample i, we restrict the predicted and reference density fields to the evaluation voxels and normalize each field to unit mass,

$$\tilde{\rho}_i(v) = \frac{\rho_i(v)}{\sum_u \rho_i(u) + \epsilon}, \qquad \tilde{\hat{\rho}}_i(v) = \frac{\hat{\rho}_i(v)}{\sum_u \hat{\rho}_i(u) + \epsilon}. \tag{56}$$

The per-sample Tanimoto similarity is

$$T_i = \frac{\sum_v \tilde{\rho}_i(v)\tilde{\hat{\rho}}_i(v)}{\sum_v \tilde{\rho}_i(v)^2 + \sum_v \tilde{\hat{\rho}}_i(v)^2 - \sum_v \tilde{\rho}_i(v)\tilde{\hat{\rho}}_i(v) + \epsilon}. \tag{57}$$

Table 11. **Ablation study on cDFT uptake prediction.** Errors are MAE of gas uptake in cm^3(STP)/g on the held-out cDFT test set. **Density Head** is the full MADField-cDFT model that predicts $\rho(\mathbf{r})$ and integrates it to obtain uptake. **Scalar Head** uses the same backbone but directly predicts uptake. **w/o Boltzmann residual** removes the Boltzmann-residual output parameterization. **FNO baseline** replaces the Swin backbone with a Fourier neural operator baseline. **Total** denotes the sample-weighted mean over all evaluated cDFT state points.

Configuration	H_2	Ar	Kr	Xe	N_2	CH_4	CO_2	C_2H_6	C_3H_8	**Total**
Scalar Head	2.46	2.17	4.71	11.93	1.72	9.06	8.41	10.85	17.66	5.91
w/o Boltzmann residual	0.16	0.21	0.72	3.28	0.17	1.26	1.46	2.15	4.55	1.01
FNO baseline	0.24	0.66	2.98	9.74	0.44	5.31	6.44	8.03	16.07	3.66
Density Head	**0.08**	**0.11**	**0.49**	**2.87**	**0.13**	**0.80**	**1.17**	**2.01**	**4.51**	**0.82**

Table 12. **cDFT pre-training ablation for GCMC uptake prediction.** Errors are MAE of gas uptake in cm^3(STP)/g. MOF evaluates in-distribution GCMC uptake prediction on crystalline MOFs. PRAM evaluates zero-shot transfer to disordered porous materials for CH_4. **Tot.** denotes the sample-weighted mean.

Model	MOF (ID)			PRAM (OOD, CH_4)				
	CH_4	Xe	**Tot.**	PIM	HCP	Kerogen	Bhatia	**Tot.**
MADField-GCMC-scratch	0.89	3.14	1.68	10.06	10.39	22.35	102.65	13.60
MADField-GCMC	**0.33**	**1.37**	**0.58**	**0.57**	**0.40**	**1.42**	**0.67**	**0.62**

We report the mean over the test set,

$$T_\rho = \frac{1}{N}\sum_{i=1}^{N} T_i. \tag{58}$$

A value of 1 indicates identical normalized density fields, while lower values indicate weaker spatial overlap. Because each field is normalized before comparison, T_ρ measures density-shape agreement independently of total uptake magnitude.

Aggregation. Per-adsorbate and per-family columns are computed by averaging over the corresponding state points. Unless otherwise stated, **Total** denotes the sample-weighted mean over all evaluated state points in the listed group. For MOF benchmarks, this aggregates over the evaluated adsorbate-pressure-structure state points. For PRAM transfer, this aggregates over all evaluated structures across the listed material families. We additionally report mean relative error for PRAM because methane uptake values are smaller than in the high-pressure MOF benchmark.

I. Additional Experiments

This appendix provides additional analyses for the cDFT uptake benchmark, OOD density transfer, and warm-start experiment. We include controlled ablations, zero-shot density transfer to amorphous carbon, external-potential well-depth stratification, and inference-time measurements.

I.1. Ablation Study

We report controlled comparisons that isolate three design choices. The scalar-head variant tests whether field supervision is more effective than direct uptake regression. The no-Boltzmann-residual variant tests the contribution of the Boltzmann residual parameterization. The FNO baseline tests whether the Swin volumetric backbone improves the density-prediction route under comparable supervision. Table 11 reports uptake MAE on the held-out cDFT test set. We also compare direct GCMC training with cDFT-pre-trained LoRA adaptationTable 12. MADField-GCMC-scratch is trained directly on the GCMC split without cDFT pre-training. MADField-GCMC starts from MADField-cDFT and adapts to GCMC density fields with LoRA.

Table 13. **Zero-shot cDFT density transfer to amorphous carbon.** Tanimoto similarity is computed on CH_4 density fields from the JLA-Gardner amorphous-carbon set. Higher is better.

Model	CH_4 Tanimoto ↑
DeepAPD	0.418
SorbIIT	0.492
MADField-cDFT	**0.885**

Table 14. **Inference time comparison.** Times are measured over 100 samples and reported as wall-clock seconds per sample. MADField models and the density baselines are evaluated with their benchmark scripts. IsothermNet uses cached features and is included only as a forward-only scalar baseline.

Model	Evaluation mode	sec/sample ↓	samples/sec ↑
MADField-cDFT	end-to-end eval	0.253	3.95
MADField-GCMC LoRA	end-to-end eval	0.097	10.26
SorbIIT	end-to-end eval	0.197	5.08
DeepAPD	full-grid probe eval	1.872	0.53
DeepSorption	CIF parse + DS forward	0.0191	52.43
RetNet	cached voxel load/transform + RN forward	0.0096	104.44
IsothermNet	cached-feature forward-only	0.00018	5655.15

I.2. Zero-shot cDFT Density Transfer to Amorphous Carbon

We evaluate zero-shot density-field transfer on the JLA-Gardner amorphous-carbon set, shown in Table 13. Models are trained on crystalline MOF cDFT labels and evaluated on CH_4 density fields without fine-tuning. Density quality is measured by Tanimoto similarity between normalized predicted and reference density fields, so the metric evaluates spatial adsorption-pattern agreement.

MADField-cDFT transfers substantially better than the density-field baselines, indicating that the learned cDFT density operator generalizes beyond crystalline MOFs to disordered carbon geometries.

I.3. Inference Time Analysis

We measure inference time on 100 held-out samples using the same evaluation scripts used for benchmark prediction, where table is shown in Table 14. The reported time is wall-clock time divided by the number of evaluated samples. For MADField models, the timing includes model loading, data loading, preprocessing, neural forward pass, and output evaluation inside the benchmark script. For IsothermNet, the timing is cached-feature forward-only and is therefore not directly comparable to end-to-end volumetric evaluation. For DeepAPD, we report full-grid probe evaluation because the model produces normalized density fields.

MADField-GCMC LoRA runs in 0.097 seconds per sample, and MADField-cDFT runs in 0.253 seconds per sample. The GCMC-adapted LoRA model is faster than the MADField cDFT model in this benchmark because the two timings are end-to-end script timings and use different checkpoint wrappers and evaluation heads. Both MADField variants are sub-second volumetric predictors. Compared with DeepAPD, MADField-GCMC LoRA is $19.2\times$ faster while predicting physically scaled density rather than normalized density. SorbIIT has similar end-to-end runtime but requires uptake information as input, so it is not an independent uptake predictor. The scalar baselines are faster when evaluated from cached or compact features, but they do not produce three-dimensional density fields and cannot be used for cDFT warm-starting.

I.4. Experiment specs

Software. We use PyTorch 2.3.1 with CUDA 12.1 and bf16 autocast (Paszke et al., 2019). The Python environment uses Python 3.11, NumPy 1.26, and SciPy 1.17. The PC-SAFT bulk equation of state uses FeOs 0.6.1 (Rehner et al., 2023). CIF parsing and structure handling use ASE (Hjorth Larsen et al., 2017). The cDFT solver is a custom PyTorch implementation for three-dimensional PC-SAFT cDFT with periodic and non-orthorhombic cells. MADField-GCMC uses published GCMC reference data from ARC-MOF and PRAM rather than newly generated GCMC trajectories. Experiment tracking uses Weights & Biases.

Hardware and training runs. MADField cDFT training is run on a single node with $8\times$ NVIDIA H200 GPUs using DDP. The effective batch size is 64, with batch size 8 per GPU and gradient accumulation 1. The released MADField checkpoint is obtained from a supervised $50{,}000$-step run followed by a supervised $30{,}000$-step resume run. The released checkpoint is selected at step $26{,}000$ of the second run. MADField-GCMC GCMC adaptation is run on the same $8\times$ H200 setup for $50{,}000$ LoRA fine-tuning steps. Inference timings are measured on a single H200 and are reported in Section I.3.